

Geo. A

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Attorney Docket No: 28967/35255A

PATENT APPLICATION TRANSMITTAL UNDER 37 C.F.R. 1.53

Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Transmitted herewith for filing is the patent application of

Inventor(s): Robert E. Ferrell, Kari Alitalo, David N. Finegold, Marika Karkkainen

Title: SCREENING AND THERAPY FOR LYMPHATIC DISORDERS INVOLVING THE
FLT4 RECEPTOR TYROSINE KINASE (VEGFR-3)

1. Type of Application

- ☐ This is a new application for a
☒ utility patent.
☐ design patent.

☒ This is a continuation-in-part application of International Patent Application No.
PCT/US99/06133, filed March 26, 1999.

2. Application Papers Enclosed

- | | |
|-------------------------------------|--|
| 1 | Specification Cover Sheet |
| 46 | Pages of Specification (excluding Claims, Abstract, Drawings & Sequence Listing) |
| 8 | Page(s) of Claims |
| 1 | Page(s) of Abstract |
| 7 | Sheet(s) of Drawings (Figs. 1 to 7) |
| <input checked="" type="checkbox"/> | Formal |
| <input type="checkbox"/> | Informal |
| 30 | Page(s) of Sequence Listing |

CERTIFICATION UNDER 37 CFR 1.10

I hereby certify that this Patent Application Transmittal and the documents referred to as enclosed therewith are being deposited with the United States Postal Service on **August 16, 1999**, in an envelope addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231 utilizing the "Express Mail Post Office to Addressee" service of the United States Postal Service under Mailing Label No. EM362732953US.


David A. Gass

08/16/99
JCS18 U.S. PTO

JCS18 U.S. PTO
09/375248
08/16/99

00376443-034620

3. Declaration or Oath

- ☐ Enclosed
 - ☐ Executed by (check all applicable boxes)
 - ☐ Inventor(s)
 - ☐ Legal representative of inventor(s)
(37 CFR 1.42 or 1.43)
 - ☐ Joint inventor or person showing a proprietary interest on behalf
of inventor who refused to sign or cannot be reached
 - ☐ The petition required by 37 CFR 1.47 and the statement
required by 37 CFR 1.47 are enclosed. See Item 5D below
for fee.
- ☒ Not enclosed - the undersigned attorney or agent is authorized to file this
application on behalf of the applicant(s). An executed declaration will
follow.

4. Additional Papers Enclosed

- ☐ Preliminary Amendment
- ☐ Information Disclosure Statement
- ☐ Declaration of Biological Deposit
- ☒ Computer readable copy of sequence listing containing nucleotide and/or
amino acid sequence
- ☐ Microfiche computer program
- ☐ Verified statement(s) claiming small entity status under 37 CFR 1.9 and 1.27
- ☐ Associate Power of Attorney
- ☐ Verified translation of a non-English patent application
- ☐ An assignment of the invention
- ☒ Return receipt postcard
- ☒ Other: Statement Under 37 C.F.R. §1.821(f)

5. **Priority Applications Under 35 USC 119**

Certified copies of applications from which priority under 35 USC 119 is claimed are listed below and

☐ are attached.

☐ will follow.

COUNTRY	APPLICATION NO.	FILED

6. **Filing Fee Calculation (37 CFR 1.16)**

A. ☒ **Utility Application**

CLAIMS AS FILED - INCLUDING PRELIMINARY AMENDMENT (IF ANY)						
			SMALL ENTITY		OTHER THAN A SMALL ENTITY	
	NO. FILED	NO. EXTRA	RATE	FEE	RATE	FEE
BASIC FEE				\$380.00		\$760.00
TOTAL	36 -20	= 16	X 9 =	\$--	X 18 =	\$288.00
INDEP.	11 - 3	= 8	X 39 =	\$--	X 78 =	\$624.00
<input type="checkbox"/> First Presentation of Multiple Dependent Claim			+ 130 =	\$--	+ 260 =	\$--
Filing Fee:				\$--	OR	\$1,672.00

B. ☐ **Design Application (\$155.00/\$310.00)** Filing Fee: \$ _____

C. ☐ **Plant Application (\$240.00/\$480.00)** Filing Fee: \$ _____

D. **Other Fees**

☐ Recording Assignment [Fee -- \$40.00 per assignment] \$ _____

☐ Petition fee for filing by other than all the inventors
or person on behalf of the inventor where inventor refused
to sign or cannot be reached [Fee -- \$130.00] \$ _____

☐ Other \$ _____

Total Fees Enclosed \$1,672.00

7. **Method of Payment of Fees**

- ☒ Enclosed check in the amount of: \$1,672.00
- ☐ Charge Deposit Account No. 13-2855 in the amount of: \$ _____
A copy of this Transmittal is enclosed.
- ☐ Not enclosed

8. **Deposit Account and Refund Authorization**

The Commissioner is hereby authorized to charge any deficiency in the amount enclosed or any additional fees which may be required during the pendency of this application under 37 CFR 1.16 or 37 CFR 1.17 or under other applicable rules (except payment of issue fees), to Deposit Account No. 13-2855. A copy of this Transmittal is enclosed.


Please refund any overpayment to Marshall, O'Toole, Gerstein, Murray & Borun at the address below.

Please direct all future communications to David A. Gass, at the address below.

Respectfully submitted,

MARSHALL, O'TOOLE, GERSTEIN,
MURRAY & BORUN
6300 Sears Tower
233 South Wacker Drive
Chicago, Illinois 60606-6402
(312) 474-6300
(312) 474-0448 (Telefacsimile)

By:


David A. Gass
Reg. No: 38,153

August 16, 1999

JOINT INVENTORS

**APPLICATION FOR
UNITED STATES LETTERS PATENT**

S P E C I F I C A T I O N

=====

TO ALL WHOM IT MAY CONCERN:

Be it known that we, **ROBERT E. FERRELL**, a citizen of the United States of America, residing at Pittsburgh, Pennsylvania, and **KARI ALITALO**, a citizen of Finland, residing at Helsinki, FINLAND, and **DAVID N. FINEGOLD**, a citizen of the United States of America, residing at Pittsburgh, Pennsylvania, and **MARIKA KARKKAINEN**, a citizen of Finland, residing at Helsinki, FINLAND have invented a new and useful **SCREENING AND THERAPY FOR LYMPHATIC DISORDERS INVOLVING THE FLT4 RECEPTOR TYROSINE KINASE (VEGFR-3)**, of which the following is a specification.

**SCREENING AND THERAPY FOR
LYMPHATIC DISORDERS INVOLVING THE
FLT4 RECEPTOR TYROSINE KINASE (VEGFR-3)**

5 This application is a Continuation-in-Part of International Patent
Application No. PCT/US99/06133, filed March 26, 1999, incorporated herein by
reference in its entirety.

ACKNOWLEDGMENT OF GOVERNMENT SUPPORT

10 This invention was made with United States and Finnish government
support, including support under contract R03-HD35174, awarded by the U.S.
National Institutes of Health. The U.S. Government has certain rights in this
invention.

FIELD OF THE INVENTION

15 The present invention relates generally to the fields of molecular
biology and medicine; more particularly to the areas of genetic screening and the
identification and treatment of hereditary disorders; and more particularly to
identification and treatment of hereditary lymphedema.

DESCRIPTION OF RELATED ART

20 The lymphatic system is a complex structure organized in parallel
fashion to the circulatory system. In contrast to the circulatory system, which utilizes
the heart to pump blood throughout the body, the lymphatic system pumps lymph fluid
using the inherent contractility of the lymphatic vessels. The lymphatic vessels are not
interconnected in the same manner as the blood vessels, but rather form a set of
coordinated structures including the initial lymphatic sinuses [Jeltsch *et al.*, *Science*,
276:1423-1425 (1997); and Castenholz, A., in Olszewski, W.L. (ed.), *Lymph Stasis:*
25 *Pathophysiology, Diagnosis, and Treatment*. CRC Press: Boca Raton, Florida (1991),
pp.15-42] which drain into the lymphatic capillaries and subsequently to the collecting
lymphatics which drain into the lymphatic trunks and the thoracic duct which
ultimately drains into the venous circulation. The composition of the channels through
which lymph passes is varied [Olszewski, W.L., in Olszewski, W.L. (ed.), *Lymph*
30 *Stasis: Pathophysiology, Diagnosis, and Treatment*. CRC Press: Boca Raton, Florida

(1991), pp. 235-258; and Kinmonth, J.B., in Kinmonth, J.B. (ed), *The Lymphatics: Diseases, Lymphography and Surgery*. Edward Arnold Publishers: London, England (1972), pp. 82-86], including the single endothelial layers of the initial lymphatics, the multiple layers of the collecting lymphatics including endothelium, muscular and adventitial layers, and the complex organization of the lymph node. The various organs of the body such as skin, lung, and GI tract have components of the lymphatics with various unique features. [See Ohkuma, M., in Olszewski (1991), *supra*, at pp. 157-190; Uhley, H. and Leeds, S., in Olszewski (1991), *supra*, at pp. 191-210; and Barrowman, J.A., in Olszewski (1991), at pp. 221-234).]

Molecular biology has identified at least a few genes and proteins postulated to have roles mediating the growth and/or embryonic development of the lymphatic system. One such gene/protein is the receptor tyrosine kinase designated Flt4 (fms-like tyrosine kinase 4), cloned from human erythroleukaemia cell and placental cDNA libraries. [See U.S. Patent No. 5,776,755; Aprelikova *et al.*, *Cancer Res.*, 52: 746-748 (1992); Galland *et al.*, *Genomics*, 13: 475-478 (1992); Galland *et al.*, *Oncogene*, 8: 1233-1240 (1993); and Pajusola *et al.*, *Cancer Res.*, 52:5738-5743 (1992), all incorporated herein by reference.] Studies showed that, in mouse embryos, a targeted disruption of the *Flt4* gene leads to a failure of the remodeling of the primary vascular network, and death after embryonic day 9.5 [Dumont *et al.*, *Science*, 282: 946-949 (1998)]. These studies suggested that *Flt4* has an essential role in the development of the embryonic vasculature, before the emergence of the lymphatic vessels. However, additional studies indicated that, during further development, the expression of *Flt4* becomes restricted mainly to lymphatic vessels [Kaipainen, *et al.*, *Proc. Natl. Acad. Sci. USA*, 92: 3566-3570 (1995)].

In humans, there are two isoforms of the Flt4 protein, designated as Flt4s (short, Genbank Accession No. X68203) and Flt4l (long, Genbank Accession Nos. X68203 and S66407, SEQ ID NO: 1). The sequence of these isoforms is largely identical, except for divergence that occurs at the carboxyl terminus of the receptor as a result of alternative mRNA splicing at the 3' end. The C-terminus of the long form contains three tyrosyl residues, and one of them (Y1337 (SEQ ID NO: 2)) serves as an autophosphorylation site in the receptor [Fournier *et al.*, *Oncogene*, 11: 921-931 (1995); and Pajusola, *et al.*, *Oncogene*, 8: 2931-2937 (1993)]. Only the long form is detected in human erythroleukaemia (HEL) and in a megakaryoblastic cell line (the

DAMI cells), and the mouse *Flt4* gene (Genbank Accession No. L07296) only produces one mRNA transcript, corresponding to Flt4l [Galland *et al.*, *Oncogene*, 8: 1233-1240 (1993); and Pajusola *et al.*, *Cancer Res.*, 52: 5738-5743 (1992)]. These findings suggest that the long form of Flt4 may be responsible for most of the biological properties of this receptor. The Flt4 protein is glycosylated and proteolytically processed in transfected cells [Pajusola *et al.*, *Oncogene*, 9: 3545-3555 (1994)]. During this process, the 175 kD form of the receptor matures to a 195 kD form, which is subsequently cleaved into a 125 kD C-terminal fragment, and a 75 kD extracellular domain-containing fragment, which are linked by disulphide bonding in the mature receptor.

Two growth factors, named vascular endothelial growth factors C and D (VEGF-C and VEGF-D) due to amino acid sequence similarity to earlier-discovered vascular endothelial growth factor, have been shown to bind and activate the tyrosine phosphorylation of Flt4. [Achen *et al.*, *Proc. Natl. Acad. Sci. USA*, 95: 548-553 (1998); Joukov *et al.*, *EMBO J.*, 16: 3898-3911; and Joukov *et al.*, *EMBO J.*, 15: 290-298 (1996)]. Because of Flt4's growth factor binding properties and the fact that Flt4 possesses amino acid sequence similarity to two previously identified VEGF receptors (Flt1/VEGFR-1 and KDR/VEGFR-2), Flt4 has also been designated VEGFR-3, and these terms are used interchangeably herein.

When VEGF-C was intentionally over-expressed under a basal keratin promoter in transgenic mice, a hyperplastic lymphatic vessel network in the skin was observed. [Jeltsch *et al.*, *Science*, 276:1423-1425 (1997).] The results of this study, when combined with the expression pattern of VEGFR-3 in the lymphatic vasculature, suggest that lymphatic growth may be induced by VEGF-C and mediated via VEGFR-3. Notwithstanding the foregoing insights involving one cell surface receptor and the two apparent ligands therefor, little is known about the developmental regulation of the lymphatic system.

Hereditary or primary lymphedema, first described by Milroy in 1892 [Milroy, *N.Y. Med. J.*, 56:505-508 (1892)], is a developmental disorder of the lymphatic system which leads to a disabling and disfiguring swelling of the extremities. Hereditary lymphedema generally shows an autosomal dominant pattern of inheritance with reduced penetrance, variable expression, and variable age-at-onset [Greenlee *et al.*, *Lymphology*, 26:156-168 (1993)]. Swelling may appear in one or all limbs,

varying in degree and distribution. If untreated, such swelling worsens over time. In rare instances, angiosarcoma may develop in affected tissues [Offori *et al.*, *Clin. Exp. Dermatol.*, 18:174-177 (1993)]. Despite having been described over a century ago, little progress has been made in understanding the mechanisms causing lymphedema.

5 A long-felt need exists for the identification of the presumed genetic variations that underlie hereditary lymphedema, to permit better informed genetic counseling in affected families, earlier diagnosis and treatment, and the development of more targeted and effective lymphedema therapeutic regimens. In addition, identification of genetic markers and high risk members of lymphedema families facilitates the
10 identification and management of environmental factors that influence the expression and severity of a lymphedema phenotype.

SUMMARY OF THE INVENTION

The present invention provides materials and methods that address one or more of the long-felt needs identified above by identifying a genetic marker that
15 correlates and is posited to have a causative role in the development of hereditary lymphedema. The invention is based in part on the discovery that, in several families with members afflicted with hereditary lymphedema, the lymphedema phenotype correlates with genetic markers localized to chromosome 5q34-q35; and that in at least some such families, a missense mutation in the VEGFR-3 gene (which maps to
20 chromosome 5q34-q35) exists that appears to behave in a loss-of-function dominant negative manner to decrease tyrosine kinase signaling of the receptor. In view of the fact that VEGFR-3 acts as a high affinity receptor for vascular endothelial growth factor C (VEGF-C), a growth factor whose effects include modulation of the growth of the lymphatic vascular network, these linkage and biochemical studies provide an
25 important marker for determining a genetic predisposition for lymphedema in healthy individuals; and for diagnosing hereditary lymphedema in symptomatic individuals. Materials and methods for performing such genetic analyses are considered aspects of the present invention.

Thus, the invention provides genetic screening procedures that entail
30 analyzing a person's genome -- in particular their *VEGFR-3* alleles -- to determine whether the individual possesses a genetic characteristic found in other individuals that are considered to be afflicted with, or at risk for, developing hereditary lymphedema.

For example, in one embodiment, the invention provides a method for determining a hereditary lymphedema development potential in a human subject comprising the steps of analyzing the coding sequence of the VEGFR-3 genes from the human subject; and determining hereditary lymphedema development potential in said human subject from the analyzing step.

In another embodiment, the invention provides a method of screening a human subject for an increased risk of developing a lymphatic disorder, comprising the steps of: (a) assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering the encoded VEGFR-3 amino acid sequence or expression of at least one VEGFR-3 allele; and (b) screening for an increased risk of developing a lymphatic disorder from the presence or absence of said mutation.

By "human subject" is meant any human being, human embryo, or human fetus. It will be apparent that methods of the present invention will be of particular interest to individuals that have themselves been diagnosed with lymphedema or have relatives that have been diagnosed with lymphedema.

By "screening for an increased risk" is meant determination of whether a genetic variation exists in the human subject that correlates with a greater likelihood of developing lymphedema than exists for the human population as a whole, or for a relevant racial or ethnic human sub-population to which the individual belongs. Both positive and negative determinations (i.e., determinations that a genetic predisposition marker is present or is absent) are intended to fall within the scope of screening methods of the invention. In preferred embodiments, the presence of a mutation altering the sequence or expression of at least one *Flt4* receptor tyrosine kinase allele in the nucleic acid is correlated with an increased risk of developing a lymphatic disorder, whereas the absence of such a mutation is reported as a negative determination.

By "lymphatic disorder" is meant any clinical condition affecting the lymphatic system, including but not limited to lymphedemas, lymphangiomas, lymphangiosarcomas, lymphangiomatosis, lymphangiectasis, and cystic hygroma.

Preferred embodiments are methods of screening a human subject for an increased risk of developing a lymphedema disorder, i.e., any disorder that physicians would diagnose as lymphedema and that is characterized by swelling associated with lymph accumulation, other than lymphedemas for which non-genetic causes (e.g., parasites,

surgery) are known. By way of example, lymphedema disorders include Milroy-Nonne (OMIM 153100) syndrome-early onset lymphedema [Milroy, *N.Y. Med. J.*, 56:505-508 (1892); and Dale, *J. Med. Genet.*, 22: 274-278 (1985)] and lymphedema praecox (Meige syndrome, OMIM 153200)-late onset lymphedema [Lewis *et al.*, *J. Ped.*, 104:641-648 (1984); Holmes *et al.*, *Pediatrics* 61:575-579 (1978); and Wheeler *et al.*, *Plastic Reconstructive Surg.*, 67:362-364 (1981)] which generally are described as separate entities, both characterized by dominant inheritance. However, there is confusion in the literature about the separation of these disorders. In Milroy's syndrome, the presence of edema, which is usually more severe in the lower extremities, is seen from birth. Lymphedema praecox presents in a similar fashion but the onset of swelling is usually around puberty. Some cases have been reported to develop in the post-pubertal period. In the particular analyses described herein, the lymphedema families showing linkage to 5q34-q35 show an early onset for most affected individuals, but individuals in these pedigrees have presented during or after puberty.

The "assaying" step of the invention may involve any techniques available for analyzing nucleic acid to determine its characteristics, including but not limited to well-known techniques such as single-strand conformation polymorphism analysis (SSCP) [Orita *et al.*, *Proc Natl. Acad. Sci. USA*, 86: 2766-2770 (1989)]; heteroduplex analysis [White *et al.*, *Genomics*, 12: 301-306 (1992)]; denaturing gradient gel electrophoresis analysis [Fischer *et al.*, *Proc. Natl. Acad. Sci. USA*, 80: 1579-1583 (1983); and Riesner *et al.*, *Electrophoresis*, 10: 377-389 (1989)]; DNA sequencing; RNase cleavage [Myers *et al.*, *Science*, 230: 1242-1246 (1985)]; chemical cleavage of mismatch techniques [Rowley *et al.*, *Genomics*, 30: 574-582 (1995); and Roberts *et al.*, *Nucl. Acids Res.*, 25: 3377-3378 (1997)]; restriction fragment length polymorphism analysis; single nucleotide primer extension analysis [Shumaker *et al.*, *Hum. Mutat.*, 7: 346-354 (1996); and Pastinen *et al.*, *Genome Res.*, 7: 606-614 (1997)]; 5' nuclease assays [Pease *et al.*, *Proc. Natl. Acad. Sci. USA*, 91:5022-5026 (1994)]; DNA Microchip analysis [Ramsay, G., *Nature Biotechnology*, 16: 40-48 (1999); and Chee *et al.*, U.S. Patent No. 5,837,832]; and ligase chain reaction [Whiteley *et al.*, U.S. Patent No. 5,521,065]. [See generally, Schafer and Hawkins, *Nature Biotechnology*, 16: 33-39 (1998).] All of the foregoing documents are hereby incorporated by reference in their entirety.

In one preferred embodiment, the assaying involves sequencing of nucleic acid to determine nucleotide sequence thereof, using any available sequencing technique. [See, e.g., Sanger *et al.*, *Proc. Natl. Acad. Sci. (USA)*, 74: 5463-5467 (1977) (dideoxy chain termination method); Mirzabekov, *TIBTECH*, 12: 27-32 (1994) (sequencing by hybridization); Drmanac *et al.*, *Nature Biotechnology*, 16: 54-58 (1998); U.S. Patent No. 5,202,231; and *Science*, 260: 1649-1652 (1993) (sequencing by hybridization); Kieleczawa *et al.*, *Science*, 258: 1787-1791 (1992) (sequencing by primer walking); (Douglas *et al.*, *Biotechniques*, 14: 824-828 (1993) (Direct sequencing of PCR products); and Akane *et al.*, *Biotechniques* 16: 238-241 (1994); Maxam and Gilbert, *Meth. Enzymol.*, 65: 499-560 (1977) (chemical termination sequencing), all incorporated herein by reference.] The analysis may entail sequencing of the entire *VEGFR-3* gene genomic DNA sequence, or portions thereof; or sequencing of the entire *VEGFR-3* coding sequence or portions thereof. In some circumstances, the analysis may involve a determination of whether an individual possesses a particular *VEGFR-3* allelic variant, in which case sequencing of only a small portion of nucleic acid -- enough to determine the sequence of a particular codon characterizing the allelic variant -- is sufficient. This approach is appropriate, for example, when assaying to determine whether one family member inherited the same allelic variant that has been previously characterized for another family member, or, more generally, whether a person's genome contains an allelic variant that has been previously characterized and correlated with heritable lymphedema. More generally, the sequencing may be focused on those portions of the *VEGFR-3* sequence that encode a VEGFR-3 kinase domain, since several different and apparently causative mutations in affected individuals that have been identified correspond to residues within an intracellular VEGFR-3 kinase domain. Referring to SEQ ID NOs: 1 and 2, the two kinase domains of human wild type VEGFR-3 correspond to nucleotides 2546 to 2848 and 3044 to 3514 of SEQ ID NO: 1, which encode residues 843 to 943 and 1009 to 1165 of SEQ ID NO: 2. Such kinase domains are localized to exons 17-20 and 22-26 in the *VEGFR-3* gene, so the sequencing/analysis may be focused on those exons in particular. Molecular modeling suggests that, within these domains, residues G852, G854, G857, K879, E896, H1035, D1037, N1042, D1055, F1056, G1057, E1084, D1096, and R1159 are of particular importance in comprising or shaping the catalytic pocket of the VEGFR-3 kinase domains, so the sequencing may focus on

these residues (in addition to residues described herein for which mutations have already been identified).

In a related embodiment, the invention provides PCR primers useful for amplifying particular exon sequences of human VEGFR-3 genomic DNA. The Examples below identify preferred primers for amplifying Exon 17, Exon 22, and Exon 24 sequences, where specific missense mutations described herein map. In addition, the Examples below describe the Exon-Intron junctions of human VEGFR-3, which, in combination with the VEGFR-3 cDNA sequence provided herein, permit the manufacture of appropriate oligonucleotide primers for other exons. Any such primers of, *e.g.*, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more nucleotides that are identical or exactly complementary to a human VEGFR-3 genomic sequence and that includes or is within 50 nucleotides of a VEGFR-3 exon-intron splice site is intended to be within the scope of the invention.

In another embodiment, the assaying step comprises performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences. In a preferred embodiment, the hybridization involves a determination of whether nucleic acid derived from the human subject will hybridize with one or more oligonucleotides, wherein the oligonucleotides have nucleotide sequences that correspond identically to a portion of the VEGFR-3 gene sequence, preferably the VEGFR-3 coding sequence set forth in SEQ ID NO: 1, or that correspond identically except for one mismatch. The hybridization conditions are selected to differentiate between perfect sequence complementarity and imperfect matches differing by one or more bases. Such hybridization experiments thereby can provide single nucleotide polymorphism sequence information about the nucleic acid from the human subject, by virtue of knowing the sequences of the oligonucleotides used in the experiments.

Several of the techniques outlined above involve an analysis wherein one performs a polynucleotide migration assay, *e.g.*, on a polyacrylamide electrophoresis gel, under denaturing or non-denaturing conditions. Nucleic acid derived from the human subject is subjected to gel electrophoresis, usually adjacent to one or more reference nucleic acids, such as reference VEGFR-3 sequences having a coding sequence identical to all or a portion of SEQ ID NO: 1, or identical except for

one known polymorphism. The nucleic acid from the human subject and the reference sequence(s) are subjected to similar chemical or enzymatic treatments and then electrophoresed under conditions whereby the polynucleotides will show a differential migration pattern, unless they contain identical sequences. [See generally Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology*, New York: John Wiley & Sons, Inc. (1987-1999); and Sambrook *et al.*, (eds.), *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989), both incorporated herein by reference in their entirety.]

In the context of assaying, the term "nucleic acid of a human subject" is intended to include nucleic acid obtained directly from the human subject (*e.g.*, DNA or RNA obtained from a biological sample such as a blood, tissue, or other cell or fluid sample); and also nucleic acid derived from nucleic acid obtained directly from the human subject. By way of non-limiting examples, well known procedures exist for creating cDNA that is complementary to RNA derived from a biological sample from a human subject, and for amplifying (*e.g.*, via polymerase chain reaction (PCR)) DNA or RNA derived from a biological sample obtained from a human subject. Any such derived polynucleotide which retains relevant nucleotide sequence information of the human subject's own DNA/RNA is intended to fall within the definition of "nucleic acid of a human subject" for the purposes of the present invention.

In the context of assaying, the term "mutation" includes addition, deletion, and/or substitution of one or more nucleotides in the *VEGFR-3* gene sequence. The invention is demonstrated by way of non-limiting examples set forth below that identify several mutations in *VEGFR-3*, including single nucleotide polymorphisms that introduce missense mutations into the *VEGFR-3* coding sequence (as compared to the *VEGFR-3* cDNA sequence set forth in SEQ ID NO: 1) and other polymorphisms that occur in introns and that are identifiable via sequencing, restriction fragment length polymorphism, or other techniques. Example 2 provides an assay to determine whether a *VEGFR-3* mutation inhibits VEGFR-3 signaling. Additional assays to study both ligand binding and signaling activities of VEGFR-3 are disclosed, *e.g.*, in U.S. Patent No. 5,776,755 and International Patent Publication No. WO 98/33917, published 06 August 1998, both of which are incorporated herein by reference in their entirety. Evidence that a VEGFR-3 mutation inhibits VEGFR-3 signaling is evidence that the mutation may have a causative role in lymphedema

phenotype. However, even mutations that have no apparent causative role may serve as useful markers for heritable lymphedema, provided that the appearance of the mutation correlates reliably with the appearance of lymphedema.

5 In a related embodiment, the invention provides a method of screening for a VEGFR-3 hereditary lymphedema genotype in a human subject, comprising the steps of: (a) providing a biological sample comprising nucleic acid from a human subject; (b) analyzing the nucleic acid for the presence of a mutation or mutations in a VEGFR-3 allele in the nucleic acid of the human subject; (c) determining a VEGFR-3 genotype from said analyzing step; and (d) correlating the presence of a mutation in a
10 VEGFR-3 allele with a hereditary lymphedema genotype. In a preferred embodiment, the biological sample is a cell sample containing human cells that contain genomic DNA of the human subject.

Although more time consuming and expensive than methods involving nucleic acid analysis, the invention also may be practiced by assaying protein of a
15 human subject to determine the presence or absence of an amino acid sequence variation in VEGFR-3 protein from the human subject. Such protein analyses may be performed, e.g., by fragmenting VEGFR-3 protein via chemical or enzymatic methods and sequencing the resultant peptides; or by Western analyses using an antibody having specificity for a particular allelic variant of VEGFR-3.

20 The invention also provides materials that are useful for performing methods of the invention. For example, the present invention provides oligonucleotides useful as probes in the many analyzing techniques described above. In general, such oligonucleotide probes comprise 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38,
25 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides that have a sequence that is identical, or exactly complementary, to a portion of a human VEGFR-3 gene sequence, or that is identical or exactly complementary except for one nucleotide substitution. In a preferred embodiment, the oligonucleotides have a sequence that corresponds in the foregoing manner to a human VEGFR-3 coding sequence, and in
30 particular, the VEGFR-3 coding sequence set forth in SEQ ID NO: 1. In one variation, an oligonucleotide probe of the invention is purified and isolated. In another variation, the oligonucleotide probe is labeled, e.g., with a radioisotope, chromophore,

or fluorophore. In yet another variation, the probe is covalently attached to a solid support. [See generally Ausubel *et al.* And Sambrook *et al.*, *supra*.]

In preferred embodiments, the invention comprises an oligonucleotide probe useful for detecting one or more of several mutations that have been

5 characterized herein in affected individuals, including:

- (1) a missense mutation at nucleotide 3360 of SEQ ID NO: 1, causing a proline to leucine change at residue 1114 in SEQ ID NO: 2;
- (2) a missense mutation at nucleotide 2588 of SEQ ID NO: 1, causing a glycine to arginine change at residue 857 in SEQ ID NO: 2;
- 10 (3) a missense mutation at nucleotide 3141 of SEQ ID NO: 1, causing an arginine to proline change at residue 1041 in SEQ ID NO: 2;
- (4) a missense mutation at nucleotide 3150 in SEQ ID NO: 1, causing a leucine to proline change at residue 1044 in SEQ ID NO: 2; and
- (5) a missense mutation at nucleotide 3164 of SEQ ID NO: 1, causing an
15 aspartic acid to asparagine change at residue 1049 in SEQ ID NO: 2.

For example, the invention provides oligonucleotides comprising anywhere from 6 to 50 nucleotides that have a sequence that is identical to, or exactly complementary to, a portion of the human VEGFR-3 coding sequence set forth in SEQ ID NO: 1, except for a nucleotide substitution corresponding to nucleotide 3360 of SEQ ID NO: 1.

20 Such oligonucleotides may be generically described by the formula X_nYZ_m or its complement; where n and m are integers from 0 to 49; where $5 \leq (n + m) \leq 49$; where X_n is a stretch of n nucleotides identical to a first portion of SEQ ID NO: 1 and Z_m is a stretch of m nucleotides identical to a second portion of SEQ ID NO: 1, wherein the first and second portions are separated in SEQ ID NO: 1 by one nucleotide; and
25 wherein Y represents a nucleotide other than the nucleotide that separates the first and second portions of SEQ ID NO: 1. For example, where X_n represents 0 to 49 nucleotides immediately upstream (5') of nucleotide 3360 of SEQ ID NO: 1 and Z_m represents 0 to 49 nucleotides immediately downstream (3') of nucleotide 3360 of SEQ ID NO: 1, Y represents a nucleotide other than cytosine, since a cytosine
30 nucleotide is found at position 3360 of SEQ ID NO: 1. In a preferred embodiment, Y is a thymine nucleotide. Similar examples are contemplated for the other specific mutations identified immediately above.

In a related embodiment, the invention provides a kit comprising at least two such oligonucleotide probes. Preferably, the two or more probes are provided in separate containers, or attached to separate solid supports, or attached separately to the same solid support, e.g., on a DNA microchip.

5 In still another related embodiment, the invention provides an array of oligonucleotide probes immobilized on a solid support, the array having at least 4 probes, preferably at least 100 probes, and preferably up to 100,000, 10,000, or 1000 probes, wherein each probe occupies a separate known site in the array. In a preferred embodiment, the array includes probe sets comprising two to four probes, wherein one
10 probe is exactly identical or exactly complementary to a human VEGFR-3 coding sequence, and the other one to three members of the set are exactly identical to the first member, but for at least one different nucleotide, which different nucleotide is located in the same position in each of the one to three additional set members. In one preferred embodiment, the array comprises several such sets of probes, wherein the
15 sets correspond to different segments of the human VEGFR-3 gene sequence. In a highly preferred embodiment, the array comprises enough sets of oligonucleotides of length N to correspond to every particular N-mer sequence of the VEGFR-3 gene, where N is preferably 6 to 25 and more preferably 9 to 20. Materials and methods for making such probes are known in the art and are described, for example, in U.S. Patent
20 Nos. 5,837,832, 5,202,231, 5,002,867, and 5,143,854.

Moreover, the discoveries which underlie the present invention identify a target for therapeutic intervention in cases of hereditary lymphedema. The causative mutations in the families that have been studied in greatest detail are mutations that appear to result in VEGFR-3 signaling that is reduced in heterozygous affected
25 individuals, but not completely eliminated. This data supports a therapeutic indication for administration of agents, such as VEGFR-3 ligand polypeptides, that will induce VEGFR-3 signaling in the lymphatic endothelia of affected individuals to effect improvement in the structure and function of the lymphatic vasculature of such individuals. In addition, therapeutic gene therapy, to replace defective VEGFR-3
30 alleles or increase production of VEGFR-3 ligand polypeptides *in vivo*, is envisioned as an aspect of the invention.

Thus, in yet another aspect, the invention provides a therapeutic or prophylactic method of treatment for lymphedema, comprising the step of

administering to a mammalian subject in need of therapeutic or prophylactic treatment for lymphedema a composition comprising a compound effective to induce intracellular signaling of VEGFR-3 in lymphatic endothelial cells that express said receptor. In a preferred embodiment, the compound comprises a polypeptide ligand for VEGFR-3, or a polynucleotide encoding such a ligand, wherein the polynucleotide is administered in a form that results in transcription and translation of the polynucleotide in the mammalian subject to produce the ligand *in vivo*. In another preferred embodiment, the compound comprises any small molecule that is capable of binding to the VEGFR-3 receptor extracellular or intracellular domain and inducing intracellular signaling.

For example, the invention provides a therapeutic or prophylactic method of treatment for lymphedema, comprising the step of administering to a mammalian subject in need of therapeutic or prophylactic treatment for lymphedema a composition comprising a polynucleotide, the polynucleotide comprising a nucleotide sequence that encodes a vascular endothelial growth factor C (VEGF-C) polypeptide. In a preferred embodiment, the subject is a human subject.

While it is contemplated that the VEGF-C polynucleotide could be administered purely as a prophylactic treatment to prevent lymphedema in subjects at risk for developing lymphedema, it is contemplated in a preferred embodiment that the polynucleotide be administered to subjects afflicted with lymphedema, for the purpose of ameliorating its symptoms (e.g., swelling due to the accumulation of lymph). The polynucleotide is included in the composition in an amount and in a form effective to promote expression of a VEGF-C polypeptide in or near the lymphatic endothelia of the mammalian subject, to stimulate VEGFR-3 signaling in the lymphatic endothelia of the subject.

In a preferred embodiment, the mammalian subject is a human subject. Practice of methods of the invention in other mammalian subjects, especially mammals that are conventionally used as models for demonstrating therapeutic efficacy in humans (e.g., primate, porcine, canine, equine, murine, or rabbit animals), also is contemplated. Several potential animal models for hereditary lymphedema have been described in the literature. [See, e.g., Lyon *et al.*, *Mouse News Lett.* 71: 26 (1984), *Mouse News Lett.* 74: 96 (1986), and *Genetic variants and strains of the laboratory mouse*, 2nd ed., New York: Oxford University Press (1989), p. 70 (*Chylous ascites mouse*); Dumont *et al.*, *Science*, 282: 946-949 (1998) (heterozygous VEGFR-3

knockout mouse); Patterson *et al.*, "Hereditary Lymphedema," *Comparative Pathology Bulletin*, 3: 2 (1971) (canine hereditary lymphedema model); van der Putte, "Congenital Hereditary Lymphedema in the Pig," *Lympho*, 11: 1-9 (1978); and Campbell-Beggs *et al.*, "Chyloabdomen in a neonatal foal," *Veterinary Record*, 137: 96-98 (1995).] Those models which are determined to have analogous mutations to the VEGFR-3 gene, such as the *Chylous ascetei* (*Chy*) mouse, are preferred. The present inventors have analyzed the *VEGFR-3* genes of the *Chy* mouse and determined that affected mice contain a missense mutation that results in a phenylalanine (rather than an isoleucine) in the VEGFR-3 sequence at a position corresponding to the isoleucine at position 1053 of SEQ ID NO: 2. This mutation maps to the catalytic pocket region of the tyrosine kinase domain of the VEGFR-3 protein, and may represent a viable model for identical mutations in human (if discovered) or other mutations in humans that similarly affect the tyrosine kinase catalytic domain. The *Chy* mouse has peripheral swelling (oedema) after birth and chyle ascites. In another embodiment, "knock in" homologous recombination genetic engineering strategies are used to create an animal model (e.g., a mouse model) having a VEGFR-3 allelic variation analogous to the human variations described herein. [See, e.g., Partanen *et al.*, *Genes & Development*, 12: 2332-2344 (1998) (gene targeting to introduce mutations into a receptor protein (FGFR-1) in mice).] Such mice can also be bred to the heterozygous VEGFR-3 knockout mice or *Chy* mice described above to further modify the phenotypic severity of the lymphedema disease.

For the practice of methods of the invention, the term "VEGF-C polypeptide" is intended to include any polypeptide that has a VEGF-C or VEGF-C analog amino acid sequence (as defined elsewhere herein in greater detail) and that is able to bind the VEGFR-3 extracellular domain and stimulate VEGFR-3 signaling *in vivo*. The term "VEGF-C polynucleotide" is intended to include any polynucleotide (e.g., DNA or RNA, single- or double-stranded) comprising a nucleotide sequence that encodes a VEGF-C polypeptide. Due to the well-known degeneracy of the genetic code, multiple VEGF-C polynucleotide sequences exist that encode any selected VEGF-C polypeptide. Preferred VEGF-C polynucleotides, polypeptides, and VEGF-C variants and analogs for use in this invention are disclosed in International Patent Application No. PCT/US98/01973, published as WO 98/33917, incorporated herein by reference in its entirety.

For treatment of humans, VEGF-C polypeptides with an amino acid sequence of a human VEGF-C are highly preferred, and polynucleotides comprising a nucleotide sequence of a human VEGF-C cDNA are highly preferred. By "human VEGF-C" is meant a polypeptide corresponding to a naturally occurring protein (prepro-protein, partially-processed protein, or fully-processed mature protein) encoded by any allele of the human VEGF-C gene, or a polypeptide comprising a biologically active fragment of a naturally-occurring mature protein. By way of example, a human VEGF-C comprises a continuous portion of the amino acid sequence set forth in SEQ ID NO: 4 sufficient to permit the polypeptide to bind and stimulate VEGFR-3 phosphorylation in cells that express such receptors. A polypeptide comprising amino acids 131-211 of SEQ ID NO: 4 is specifically contemplated. For example, polypeptides having an amino acid sequence comprising a continuous portion of SEQ ID NO: 4, the continuous portion having, as its amino terminus, an amino acid selected from the group consisting of positions 30-131 of SEQ ID NO: 4, and having, as its carboxyl terminus, an amino acid selected from the group consisting of positions 211-419 of SEQ ID NO: 4 are contemplated. An amino terminus selected from the group consisting of positions 102-131 of SEQ ID NO: 4 is preferred, and an amino terminus selected from the group consisting of positions 103-113 of SEQ ID NO: 4 is highly preferred. Likewise, a carboxyl terminus selected from the group consisting of positions 211-227 of SEQ ID NO: 4 is preferred. As stated above, the term "human VEGF-C" also is intended to encompass polypeptides encoded by allelic variants of the human VEGF-C characterized by the sequences set forth in SEQ ID NOs: 3 & 4.

Moreover, since the therapeutic VEGF-C is to be administered as recombinant VEGF-C or indirectly via somatic gene therapy, it is within the skill in the art to make and use analogs of human VEGF-C (and polynucleotides that encode such analogs) wherein one or more amino acids have been added, deleted, or replaced with other amino acids, especially with conservative replacements, and wherein the VEGFR-3-stimulatory biological activity has been retained. Analogs that retain VEGFR-3-stimulatory VEGF-C biological activity are contemplated as VEGF-C polypeptides for use in the present invention. In a preferred embodiment, analogs having 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 such modifications and that retain VEGFR-3-stimulatory VEGF-C biological

activity are contemplated as VEGF-C polypeptides for use in the present invention. Analogs having a deletion of or substitution for the cysteine residue at position 156 of SEQ ID NO: 4 and that retain VEGFR-3 stimulatory activity but have reduced activity toward the receptor VEGFR-2, which is expressed in blood vessels, are specifically contemplated. See WO 98/33917. Polynucleotides encoding such analogs are generated using conventional PCR, site-directed mutagenesis, and chemical synthesis techniques.

Also contemplated as VEGF-C polypeptides are non-human mammalian or avian VEGF-C polypeptides and polynucleotides. By "mammalian VEGF-C" is meant a polypeptide corresponding to a naturally occurring protein (prepro-protein, partially-processed protein, or fully-processed mature protein) encoded by any allele of a VEGF-C gene of any mammal, or a polypeptide comprising a biologically active fragment of a mature protein. The term "mammalian VEGF-C polypeptide" is intended to include analogs of mammalian VEGF-C's that possess the *in vivo* VEGFR-3-stimulatory effects of the mammalian VEGF-C.

Irrespective of which encoded VEGF-C polypeptide is chosen, any VEGF-C polynucleotide gene therapy pharmaceutical encoding it preferably comprises a nucleotide sequence encoding a secretory signal peptide fused in-frame with the VEGF-C polypeptide sequence. The secretory signal peptide directs secretion of the VEGF-C polypeptide by the cells that express the polynucleotide, and is cleaved by the cell from the secreted VEGF-C polypeptide. For example, the VEGF-C polynucleotide could encode the complete prepro-VEGF-C sequence set forth in SEQ ID NO: 4; or could encode the VEGF-C signal peptide fused in-frame to a sequence encoding a fully-processed VEGF-C (*e.g.*, amino acids 103-227 of SEQ ID NO: 4) or VEGF-C analog. Moreover, there is no requirement that the signal peptide be derived from VEGF-C. The signal peptide sequence can be that of another secreted protein, or can be a completely synthetic signal sequence effective to direct secretion in cells of the mammalian subject.

In one embodiment, the VEGF-C polynucleotide of the invention comprises a nucleotide sequence that will hybridize to a polynucleotide that is complementary to the human VEGF-C cDNA sequence specified in SEQ ID NO: 3 under the following exemplary stringent hybridization conditions: hybridization at 42°C in 50% formamide, 5X SSC, 20 mM Na₂PO₄, pH 6.8; and washing in 1X SSC at

55°C for 30 minutes; and wherein the nucleotide sequence encodes a polypeptide that binds and stimulates human VEGFR-3. It is understood that variation in these exemplary conditions occur based on the length and GC nucleotide content of the sequences to be hybridized. Formulas standard in the art are appropriate for determining appropriate hybridization conditions. [See Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (Second ed., Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 1989) §§ 9.47-9.51.]

In preferred embodiments, the VEGF-C polynucleotide further comprises additional sequences to facilitate the VEGF-C gene therapy. In one embodiment, a "naked" VEGF-C transgene (i.e., a transgene without a viral, liposomal, or other vector to facilitate transfection) is employed for gene therapy. In this embodiment, the VEGF-C polynucleotide preferably comprises a suitable promoter and/or enhancer sequence (*e.g.*, cytomegalovirus promoter/enhancer [Lehner *et al.*, *J. Clin. Microbiol.*, 29:2494-2502 (1991); Boshart *et al.*, *Cell*, 41:521-530 (1985)]; Rous sarcoma virus promoter [Davis *et al.*, *Hum. Gene Ther.*, 4:151 (1993)]; Tie promoter [Korhonen *et al.*, *Blood*, 86(5): 1828-1835 (1995)]; or simian virus 40 promoter) for expression in the target mammalian cells, the promoter being operatively linked upstream (i.e., 5') of the VEGF-C coding sequence. The VEGF-C polynucleotide also preferably further includes a suitable polyadenylation sequence (*e.g.*, the SV40 or human growth hormone gene polyadenylation sequence) operably linked downstream (i.e., 3') of the VEGF-C coding sequence. The polynucleotide may further optionally comprise sequences whose only intended function is to facilitate large-scale production of the vector, *e.g.*, in bacteria, such as a bacterial origin of replication and a sequence encoding a selectable marker. However, in a preferred embodiment, such extraneous sequences are at least partially cleaved off prior to administration to humans according to methods of the invention. One can manufacture and administer such polynucleotides to achieve successful gene therapy using procedures that have been described in the literature for other transgenes. See, *e.g.*, Isner *et al.*, *Circulation*, 91: 2687-2692 (1995); and Isner *et al.*, *Human Gene Therapy*, 7: 989-1011 (1996); incorporated herein by reference in the entirety.

Any suitable vector may be used to introduce the VEGF-C transgene into the host. Exemplary vectors that have been described in the literature include replication-deficient retroviral vectors, including but not limited to lentivirus vectors

[Kim *et al.*, *J. Virol.*, 72(1): 811-816 (1998); Kingsman & Johnson, *Scrip Magazine*, October, 1998, pp. 43-46.]; adeno-associated viral vectors [Gnatenko *et al.*, *J. Investig. Med.*, 45: 87-98 (1997)]; adenoviral vectors [See, *e.g.*, U.S. Patent No. 5,792,453; Quantin *et al.*, *Proc. Natl. Acad. Sci. USA*, 89: 2581-2584 (1992);
5 Stratford-Perricadet *et al.*, *J. Clin. Invest.*, 90: 626-630 (1992); and Rosenfeld *et al.*, *Cell*, 68: 143-155 (1992)]; Lipofectin-mediated gene transfer (BRL); liposomal vectors [See, *e.g.*, U.S. Patent No. 5,631,237 (Liposomes comprising Sendai virus proteins)] ; and combinations thereof. All of the foregoing documents are incorporated herein by reference in the entirety. Replication-deficient adenoviral vectors constitute a
10 preferred embodiment.

In embodiments employing a viral vector, preferred polynucleotides still include a suitable promoter and polyadenylation sequence as described above. Moreover, it will be readily apparent that, in these embodiments, the polynucleotide further includes vector polynucleotide sequences (*e.g.*, adenoviral polynucleotide
15 sequences) operably connected to the sequence encoding a VEGF-C polypeptide.

Thus, in one embodiment the composition to be administered comprises a vector, wherein the vector comprises the VEGF-C polynucleotide. In a preferred embodiment, the vector is an adenovirus vector. In a highly preferred embodiment, the adenovirus vector is replication-deficient, *i.e.*, it cannot replicate in the mammalian
20 subject due to deletion of essential viral-replication sequences from the adenoviral genome. For example, the inventors contemplate a method wherein the vector comprises a replication-deficient adenovirus, the adenovirus comprising the VEGF-C polynucleotide operably connected to a promoter and flanked on either end by adenoviral polynucleotide sequences.

25 The composition to be administered according to methods of the invention preferably comprises (in addition to the polynucleotide or vector) a pharmaceutically-acceptable carrier solution such as water, saline, phosphate-buffered saline, glucose, or other carriers conventionally used to deliver therapeutics intravascularly. Multi-gene therapy is also contemplated, in which case the
30 composition optionally comprises both the VEGF-C polynucleotide/vector and another polynucleotide/vector. As described in greater detail below, a VEGF-D transgene is a preferred candidate for co-administration with the VEGF-C transgene.

The "administering" that is performed according to the present method may be performed using any medically-accepted means for introducing a therapeutic directly or indirectly into a mammalian subject to reach the lymph or the lymphatic system, including but not limited to injections; oral ingestion; intranasal or topical administration; and the like. In a preferred embodiment, administration of the composition comprising the VEGF-C polynucleotide is performed intravascularly, such as by intravenous or intra-arterial injection, or by subcutaneous injection or local depot administration. In a highly preferred embodiment, the composition is administered locally, *e.g.*, to the site of swelling.

In still another variation, endothelial cells or endothelial progenitor cells are transfected *ex vivo* with a wild type VEGFR-3 transgene, and the transfected cells are administered to the mammalian subject.

In another aspect, the invention provides a therapeutic or prophylactic method of treating for lymphedema, comprising the step of administering to a mammalian subject in need of treatment for lymphedema a composition comprising a VEGF-C polypeptide, in an amount effective to treat or prevent swelling associated with lymphedema. Administration via one or more intravenous or subcutaneous injections is contemplated. Co-administration of VEGF-C polynucleotides and VEGF-C polypeptides is also contemplated.

In yet another embodiment, the invention provides the use of a VEGF-C polynucleotide or VEGF-C polypeptide for the manufacture of a medicament for the treatment or prevention of lymphedema.

In still another embodiment, the invention provides a therapeutic or prophylactic method of treatment for lymphedema, comprising the step of administering to a mammalian subject in need of therapeutic or prophylactic treatment of lymphedema a composition comprising a polynucleotide, the polynucleotide comprising a nucleotide sequence that encodes a vascular endothelial growth factor D (VEGF-D) polypeptide. Such methods are practiced essentially as described herein with respect to VEGF-C-encoding polynucleotides, except that polynucleotides encoding VEGF-D are employed. A detailed description of the human VEGF-D gene and protein are provided in Achen, *et al.*, *Proc. Nat'l Acad. Sci. U.S.A.*, 95(2): 548-553 (1998); International Patent Publication No. WO 98/07832, published 26 February 1998; and in Genbank Accession No. AJ000185, all incorporated herein by reference.

A cDNA and deduced amino acid sequence for prepro-VEGF-D is set forth herein in SEQ ID NOs: 5 and 6. Of course, due to the well-known degeneracy of the genetic code, multiple VEGF-D encoding polynucleotide sequence exist, any of which may be employed according to the methods taught herein.

5 As described herein in detail with respect to VEGF-C, the use of polynucleotides that encode VEGF-D fragments, VEGF-D analogs, VEGF-D allelic and interspecies variants, and the like which possess *in vivo* stimulatory effects of human VEGF-D are all contemplated as being encompassed by the present invention.

10 In yet another embodiment, the invention provides a therapeutic or prophylactic method of treatment for lymphedema, comprising the step of administering to a mammalian subject in need of treatment for lymphedema a composition comprising a VEGF-D polypeptide, in an amount effective to treat or prevent swelling associated with lymphedema. Administration via one or more intravenous or subcutaneous injections is contemplated.

15 The VEGFR-3 allelic variant polynucleotides and polypeptides described herein that were discovered and characterized by the present inventors are themselves considered aspects of the invention. Such polynucleotides and polypeptides are useful, for example, in screening assays (*e.g.*, cell-based assays or assays involving transgenic mice that express the polynucleotide in lieu of a native
20 VEGF-3 allele) to study the biological activities of VEGFR-3 variant alleles and identify compounds that are capable of modulating that activity, *e.g.*, to identify therapeutic candidates for treatment of lymphedema. Such screening assays are also considered aspects of the invention.

25 The polypeptides of the invention are intended to include complete VEGFR-3 polypeptides with signal peptide (*e.g.*, approximately residues 1 to 20 of SEQ ID NO: 2), mature VEGFR-3 polypeptides lacking any signal peptide, and recombinant variants wherein a foreign or synthetic signal peptide has been fused to the mature VEGFR-3 polypeptide. Polynucleotides of the invention include all polynucleotides that encode all such polypeptides. It will be understood that for
30 essentially any polypeptide, many polynucleotides can be constructed that encode the polypeptide by virtue of the well known degeneracy of the genetic code. All such polynucleotides are intended as aspects of the invention.

Thus, in yet another aspect, the invention provides a purified polynucleotide comprising a nucleotide sequence encoding a human VEGFR-3 protein variant, wherein said polynucleotide is capable of hybridizing to the complement of SEQ ID NO: 1 under stringent hybridization conditions, and wherein the encoded VEGFR-3 protein variant has an amino acid sequence that differs at position 1114, 857, 1041, 1044 or 1049 from the amino acid sequence set forth in SEQ ID NO: 1. Exemplary conditions are as follows: hybridization at 42°C in 50% formamide, 5X SSC, 20 mM Na•PO₄, pH 6.8; and washing in 0.2X SSC at 55°C. It is understood by those of skill in the art that variation in these conditions occurs based on the length and GC nucleotide content of the sequences to be hybridized. Formulas standard in the art are appropriate for determining appropriate hybridization conditions. [See Sambrook et al. (1989), *supra*, §§ 9.47-9.51.]

In a related embodiment, the invention provides a purified polynucleotide comprising a nucleotide sequence encoding a VEGFR-3 protein of a human that is affected with heritable lymphedema or other lymphatic disorder; wherein the polynucleotide is capable of hybridizing to the complement of SEQ ID NO: 1 under stringent hybridization conditions, and wherein the encoded polynucleotide has an amino acid sequence that differs from SEQ ID NO: 1 at at least one codon. It will be understood that conventional recombinant techniques can be used to isolate such polynucleotides from individuals affected with heritable lymphedema or their relatives. The wildtype VEGFR-3 cDNA sequence set forth in SEQ ID NO: 1 (or its complement, or fragments thereof) is used as a probe to identify and isolate VEGFR-3 sequences from nucleic acid derived from the individuals. Alternatively, PCR amplification primers based on the wildtype VEGFR-3 sequence are generated and used to amplify either VEGFR-3 genomic DNA or VEGFR-3 mRNA from the human subject. The resultant amplified genomic DNA or cDNA is sequenced to determine the variations that characterize the VEGFR-3 lymphedema allele of the individual. Preferred VEGFR-3 lymphedema alleles include, but are not limited to the P1114L, G857R, R1041P, L1044P and D1049N alleles described in detail herein.

In addition, the invention provides vectors that comprise the polynucleotides of the invention. Such vectors are useful for amplifying and expressing the VEGFR-3 proteins encoded by the polynucleotides, and for creating recombinant host cells and/or transgenic animals that express the polynucleotides. The

invention further provides a host cell transformed or transfected with polynucleotides (including vectors) of the invention. In a preferred embodiment, the host cell expresses the encoded VEGFR-3 protein on its surface. Such host cells are useful in cell-based screening assays for identifying modulators that stimulate or inhibit signaling of the encoded VEGFR-3. Modulators that stimulate VEGFR-3 signaling have utility as therapeutics to treat lymphedemas, whereas modulators that are inhibitory have utility for treating hyperplastic lymphatic conditions mediated by the allelic variant VEGFR-3. In a preferred embodiment, host cells of the invention are co-transfected with both a wildtype and an allelic variant VEGFR-3 polynucleotide, such that the cells express both receptor types on their surface. Such host cells are preferred for simulating a heterozygous VEGFR-3 genotype of many individuals affected with lymphedema.

In yet another aspect, the invention provides a transgenic mammal, e.g., mouse, characterized by a non-native VEGFR-3 allele that has been introduced into the mouse, and the transgenic progeny thereof. Preferred allelic variants include allelic variants that correlate with hereditary lymphedema in human subjects, such as an allelic variant wherein a P1114L, G857R, R1041P, L1044P or D1049N missense mutation has been introduced into the murine VEGFR-3 gene, or wherein the human P1114L, G857R, R1041P, L1044P or D1049N allelic variant has been substituted for a murine VEGFR-3 allele. Such mice are produced using standard methods. [See, e.g., Hogan *et al.* (eds.), *Manipulating the Mouse Embryo*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory (1986).] The introduction of the human-like mutations into non-human sequences is readily achieved with standard techniques, such as site-directed mutagenesis. The determination of which residues in a non-human sequence to alter to mimic the foregoing human mutations is routine since the foregoing mutations all occur in regions of the VEGFR-3 sequence that contain residues that are highly conserved between species. See Figs. 3A-3B.

In yet another aspect, the invention provides assays for identifying modulators of VEGFR-3 signaling, particularly modulators of the signaling of allelic variants of VEGFR-3 that correlate with lymphatic disorders such as heritable lymphedema. For example, the invention provides a method for identifying a modulator of intracellular VEGFR-3 signaling, comprising the steps of: contacting a cell expressing at least one mutant mammalian VEGFR-3 polypeptide in the presence and in the absence of a putative modulator compound; b) detecting VEGFR-3

signaling in the cell; and c) identifying a putative modulator compound in view of decreased or increased signaling in the presence of the putative modulator, as compared to signaling in the absence of the putative modulator.

By "mutant mammalian VEGFR-3 polypeptide" is meant a VEGFR-3 polypeptide that varies from a wildtype mammalian VEGFR-3 polypeptide (e.g., by virtue of one or more amino acid additions, deletions, or substitutions), wherein the variation is reflective of a naturally occurring variation that has been correlated with a lymphatic disorder, such as lymphedema. By way of example, the previously described substitution variations of human VEGFR-3, such as P1114L, have been correlated with heritable lymphedema. Any of the human allelic variants described above, or analogous human allelic variants having a different substitution at the indicated amino acid positions, or a non-human VEGFR-3 into which a mutation at the position corresponding to any of the described positions has been introduced are all examples of mutant mammalian VEGFR-3 polypeptides.

The detecting step can entail the detection of any parameter indicative of VEGFR-3 signaling. For example, the detecting step can entail a measurement of VEGFR-3 autophosphorylation, or a measurement of VEGFR-3-mediated cell growth, or a measurement of any step in the VEGFR-3 signaling cascade between VEGFR-3 autophosphorylation and cell growth.

In a preferred embodiment, the method is practiced with a cell that expresses the mutant mammalian VEGFR-3 polypeptide and a wildtype mammalian VEGFR-3 polypeptide. Such cells are thought to better mimic the conditions in heterozygous individuals suffering from a VEGFR-3-mediated lymphatic disorder. In a highly preferred embodiment, the mutant and wildtype VEGFR-3 polypeptides are human. In the preferred embodiments, the mutant VEGFR-3 polypeptide comprises a leucine amino acid at the position corresponding to position 1114 of SEQ ID NO: 2; an arginine at the position corresponding to position 857 of SEQ ID NO: 2; a proline amino acid at the position corresponding to position 1041 of SEQ ID NO: 2; a proline amino acid at the position corresponding to position 1044 of SEQ ID NO: 2; or an asparagine at the position corresponding to position 1049 of SEQ ID NO: 2.

Additional features and variations of the invention will be apparent to those skilled in the art from the entirety of this application, including the drawing and detailed description, and all such features are intended as aspects of the invention.

Likewise, features of the invention described herein can be re-combined into additional embodiments that are also intended as aspects of the invention, irrespective of whether the combination of features is specifically mentioned above as an aspect or embodiment of the invention. Also, only such limitations which are described herein as critical to the invention should be viewed as such; variations of the invention lacking limitations which have not been described herein as critical are intended as aspects of the invention.

In addition to the foregoing, the invention includes, as an additional aspect, all embodiments of the invention narrower in scope in any way than the variations specifically mentioned above. Although the applicant(s) invented the full scope of the claims appended hereto, the claims appended hereto are not intended to encompass within their scope the prior art work of others. Therefore, in the event that statutory prior art within the scope of a claim is brought to the attention of the applicants by a Patent Office or other entity or individual, the applicant(s) reserve the right to exercise amendment rights under applicable patent laws to redefine the subject matter of such a claim to specifically exclude such statutory prior art or obvious variations of statutory prior art from the scope of such a claim. Variations of the invention defined by such amended claims also are intended as aspects of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1F depict pedigrees of six hereditary lymphedema families (Families 101, 106, 111, 135, 105, and 127, respectively) informative for linkage. Filled symbols represent individuals with clinically documented lymphedema. Crossed symbols represent individuals with an ambiguous phenotype. An ambiguous phenotype is defined as self-reported swelling of the limbs with no known cause, without a clinical diagnosis of lymphedema. Individuals of ambiguous phenotype were coded as disease status unknown for the linkage analysis. The proband in each family is indicated by an arrow.

Figure 2 is a graph summarizing VITESSE analysis of lymphedema families with markers localized to chromosome 5q34-q35. In the graph, filled circles represent analyses for Families 101, 105, 106, and 111; open boxes represent analyses for Families 101, 106, and 111; open circles represent the VEGFR-3 gene; and open triangles represent Family 135. The one LOD confidence interval lies completely

within the interval flanked by markers D5S1353 and D5S408 and overlaps the most likely location of *Flt4* (*VEGFR-3*). Linkage is excluded for the entire region for family 135.

Figure 3A-3B depict an alignment of portions of the human (top line, SEQ ID NO: 2) and murine (bottom line, GenBank Acc. No. P35917, SEQ ID NO: 19) VEGFR-3 amino acid sequences to demonstrate similarity. Identical residues are marked with a line, and highly conserved and less conserved differences are marked with two dots or a single dot, respectively. The location of various mutations that have been observed to correlate with a heritable lymphedema phenotype are indicated immediately beneath the aligned sequences.

DETAILED DESCRIPTION OF THE INVENTION

Certain therapeutic aspects of the present invention involve the administration of Vascular Endothelial Growth Factor C or D polynucleotides and polypeptides. The growth factor VEGF-C, as well as native human, non-human mammalian, and avian polynucleotide sequences encoding VEGF-C, and VEGF-C variants and analogs, have been described in detail in International Patent Application Number PCT/US98/01973, filed 02 February 1998 and published on 06 August 1998 as International Publication Number WO 98/33917; in Joukov *et al.*, *J. Biol. Chem.*, 273(12): 6599-6602 (1998); and in Joukov *et al.*, *EMBO J.*, 16(13): 3898-3911 (1997), all of which are incorporated herein by reference in the entirety. As explained therein in detail, human VEGF-C is initially produced in human cells as a prepro-VEGF-C polypeptide of 419 amino acids. A cDNA and deduced amino acid sequence for human prepro-VEGF-C are set forth in SEQ ID NOs: 3 and 4, respectively, and a cDNA encoding human VEGF-C has been deposited with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 (USA), pursuant to the provisions of the Budapest Treaty (Deposit date of 24 July 1995 and ATCC Accession Number 97231). VEGF-C sequences from other species have also been reported. See Genbank Accession Nos. MMU73620 (*Mus musculus*); and CCY15837 (*Coturnix coturnix*) for example, incorporated herein by reference.

The prepro-VEGF-C polypeptide is processed in multiple stages to produce a mature and most active VEGF-C polypeptide of about 21-23 kD (as

assessed by SDS-PAGE under reducing conditions). Such processing includes cleavage of a signal peptide (SEQ ID NO: 4, residues 1-31); cleavage of a carboxyl-terminal peptide (corresponding approximately to amino acids 228-419 of SEQ ID NO: 4 and having a pattern of spaced cysteine residues reminiscent of a Balbiani ring 3 protein (BR3P) sequence [Dignam *et al.*, *Gene*, 88:133-40 (1990); Paulsson *et al.*, *J. Mol. Biol.*, 211:331-49 (1990)]) to produce a partially-processed form of about 29 kD; and cleavage (apparently extracellularly) of an amino-terminal peptide (corresponding approximately to amino acids 32-103 of SEQ ID NO: 4) to produce a fully-processed mature form of about 21-23 kD. Experimental evidence demonstrates that partially-processed forms of VEGF-C (*e.g.*, the 29 kD form) are able to bind the VEGFR-3 receptor, whereas high affinity binding to VEGFR-2 occurs only with the fully processed forms of VEGF-C.

Moreover, it has been demonstrated that amino acids 103-227 of SEQ ID NO: 4 are not all critical for maintaining VEGF-C functions. A polypeptide consisting of amino acids 113-213 (and lacking residues 103-112 and 214-227) of SEQ ID NO: 2 retains the ability to bind and stimulate VEGFR-3, and it is expected that a polypeptide spanning from about residue 131 to about residue 211 will retain VEGF-C biological activity. The cysteine residue at position 156 has been shown to be important for VEGFR-2 binding ability. However, VEGF-C ΔC_{156} polypeptides (*i.e.*, analogs that lack this cysteine due to deletion or substitution) remain potent activators of VEGFR-3, and are therefore considered to be among the preferred candidates for treatment of lymphedema. (It has been shown that a VEGF-C C156S serine substitution analog promotes lymphatic growth when over-expressed in the skin of transgenic mice behind the K14 promoter, in a manner analogous to what was described in Jeltsch *et al.*, *Science*, 276:1423 (1997), incorporated herein by reference.) The cysteine at position 165 of SEQ ID NO: 4 is essential for binding to either receptor, whereas analogs lacking the cysteines at positions 83 or 137 compete with native VEGF-C for binding with both receptors and are able to stimulate both receptors.

An alignment of human VEGF-C with VEGF-C from other species (performed using any generally accepted alignment algorithm) suggests additional residues wherein modifications can be introduced (*e.g.*, insertions, substitutions, and/or deletions) without destroying VEGF-C biological activity. Any position at which

aligned VEGF-C polypeptides of two or more species have different amino acids, especially different amino acids with side chains of different chemical character, is a likely position susceptible to modification without concomitant elimination of function. An exemplary alignment of human, murine, and quail VEGF-C is set forth in Figure 5 of PCT/US98/01973.

Apart from the foregoing considerations, it will be understood that innumerable conservative amino acid substitutions can be performed to a wildtype VEGF-C sequence which are likely to result in a polypeptide that retains VEGF-C biological activities, especially if the number of such substitutions is small. By “conservative amino acid substitution” is meant substitution of an amino acid with an amino acid having a side chain of a similar chemical character. Similar amino acids for making conservative substitutions include those having an acidic side chain (glutamic acid, aspartic acid); a basic side chain (arginine, lysine, histidine); a polar amide side chain (glutamine, asparagine); a hydrophobic, aliphatic side chain (leucine, isoleucine, valine, alanine, glycine); an aromatic side chain (phenylalanine, tryptophan, tyrosine); a small side chain (glycine, alanine, serine, threonine, methionine); or an aliphatic hydroxyl side chain (serine, threonine). Addition or deletion of one or a few internal amino acids without destroying VEGF-C biological activities also is contemplated.

Without intending to be limited to a particular theory, the mechanism behind the efficacy of VEGF-C in treating or preventing lymphedema is believed to relate to the ability of VEGF-C to stimulate VEGFR-3 signaling. Administration of VEGF-C in quantities exceeding those usually found in interstitial fluids is expected to stimulate VEGFR-3 in human subjects who, by virtue of a dominant negative heterozygous mutation, have insufficient VEGFR-3 signaling.

The growth factor named Vascular Endothelial Growth Factor D (VEGF-D), as well as human sequences encoding VEGF-D, and VEGF-D variants and analogs, have been described in detail in International Patent Application Number PCT/US97/14696, filed 21 August 1997 and published on 26 February 1998 as International Publication Number WO 98/07832; and in Achen, *et al.*, *Proc. Nat’l Acad. Sci. U.S.A.*, 95(2): 548-553 (1998), both incorporated herein by reference in the entirety. As explained therein in detail, human VEGF-D is initially produced in human cells as a prepro-VEGF-D polypeptide of 354 amino acids. A cDNA and deduced amino acid sequence for human prepro-VEGF-D are set forth in SEQ ID Nos: 5 and 6,

respectively. VEGF-D sequences from other species also have been reported. See Genbank Accession Nos. D89628 (*Mus musculus*); and AF014827 (*Rattus norvegicus*), for example, incorporated herein by reference.

The prepro-VEGF-D polypeptide has a putative signal peptide of 21 amino acids and is apparently proteolytically processed in a manner analogous to the processing of prepro-VEGF-C. A "recombinantly matured" VEGF-D lacking residues 1-92 and 202-354 of SEQ ID NO: 6 retains the ability to activate receptors VEGFR-2 and VEGFR-3, and appears to associate as non-covalently linked dimers. Thus, preferred VEGF-D polynucleotides include those polynucleotides that comprise a nucleotide sequence encoding amino acids 93-201 of SEQ ID NO: 6.

The subject matter of the invention is further described and demonstrated with reference to the following examples.

EXAMPLE 1

Demonstration that hereditary lymphedema is linked to the VEGFR-3 locus

The following experiments, conducted to identify a gene or genes contributing to susceptibility to develop lymphedema, demonstrated that hereditary lymphedema correlates, in at least some families, to the chromosomal locus for the VEGFR-3 gene.

OVERVIEW

Families with inherited lymphedema were identified for the purpose of conducting a linkage and positional candidate gene analysis. Thirteen distinct families from the United States and Canada were identified through referrals from lymphedema treatment centers, lymphedema support groups, and from internet correspondence (worldwide web site at www.pitt.edu/~genetics/lymph/). The study protocol was approved by the Institutional Review Board of the University of Pittsburgh and participants gave written informed consent. All members of the families were of western European ancestry. Forty members of one family ("Family 101") were examined during a family reunion by a physiatrist experienced in lymphedema treatment. Family members were considered affected with hereditary lymphedema if they exhibited asymmetry or obvious swelling of one or both legs. Members of the other 12 families were scored as affected if they had received a medical diagnosis of lymphedema, or if there were personal and family reports of extremity swelling or

asymmetry. Medical records were obtained to verify status whenever possible. For the purpose of linkage analysis, individuals with very mild or intermittent swelling, heavyset legs, obesity, or a history of leg infections as the only symptom were considered to have indeterminate disease status.

5 In the 13 families, 105 individuals were classified as affected, with a male:female ratio of 1:2.3. The age of onset of lymphedema symptoms ranged from prenatal (diagnosed by ultrasound) to age 55. When affected by normal matings were analyzed, 76 of 191 children were affected, yielding a penetrance of 80%. First degree relatives of affected individuals were considered at risk.

10 Biological samples were obtained from members of the thirteen families to conduct the genetic analyses. DNA was isolated from the EDTA-anticoagulated whole blood by the method of Miller *et al.*, *Nucleic Acids Res.*, 16: 1215 (1998), and from cytobrush specimens using the Puregene DNA isolation kit (Gentra Systems, Minneapolis, MN). Analysis of the markers used in the genome scan were performed
15 by methods recognized in the art. [See Browman *et al.*, *Am. J. Hum. Genetic.*, 63:861-869 (1998); see also the NHLBI Mammalian Genotyping Service world-wide web sites (www.marshmed.org/genetics/methods/pcr.htm; and www.marshmed.org/genetics/methods/gel.htm).

Two-point linkage analysis was conducted using an autosomal
20 dominant model predicting 80% penetrance in the heterozygous state, 99% penetrance in the homozygous state, and a 1% phenocopy rate. The frequency of the disease allele was set at 1/10,000. Microsatellite marker allele frequencies were calculated by counting founder alleles, with the addition of counts of non-transmitted alleles. Multipoint analysis was carried out using distances obtained from the Location
25 Database (LDB-<http://cedar.genetics.soton.ac.uk/public.html>). Multipoint and 2-point analyses were facilitated using the VITESSE (v1.1) program. [O'Connell, J.R. and Weeks, D.E., (1995), *Nature Genet.*, 11:402-408].

DETAILED DESCRIPTION OF METHODS AND RESULTS

30 The first family studied, Family 101, was a large, multi-generational family demonstrating early onset lymphedema. (See Fig. 1.) Forty individuals of this family were examined and DNA sampled. In addition, blood was obtained from another 11 members from mailing kits. Linkage simulation was performed using

SLINK [Weeks *et al.*, *Am. J. Hum. Genet.* 47:A204 (1990)] and linkage was analyzed using MSIM [Ott, J., *Proc. Nat. Acad. Sci. USA*, 86:4175-4178 (1989)] to estimate the potential power of two point linkage analysis in the family. Marker genotypes were simulated for a marker with heterozygosity of 0.875 under a linked ($\theta=0$) and
5 unlinked ($\theta=0.5$) model using the 51 available individuals. The simulation showed that the power to detect linkage was greater than 90% for a LOD score threshold of $Z(\theta)$ 2.0. The false positive rate was less than 5%.

Shortly thereafter, two additional families (designated Families 106 and 111) segregating for autosomal dominant lymphedema were identified. These three
10 families (Figures 1A-1C, Families 101, 106 and 111) were genotyped for 366 autosomal markers by the NHLBI Mammalian Genotyping Service (www.marshmed.org/genetics). Genotypes were checked for consistency using Pedcheck [O'Connell, J.R. and Weeks, D.E., *Am. J. Hum. Genet.*, 61:A288 (1997)]. Two point linkage analysis was performed using VITESSE [O'Connell, J.R. and
15 Weeks, D.E., *Nature Genet.*, 11:402-408 (1995)]. The model for linkage assumed an autosomal dominant model of inheritance, a disease allele frequency of 0.0001 and a penetrance of 0.80.

The results from the genomic scan can be briefly summarized as follows. A summed LOD score of greater than 4.0 was observed from distal
20 chromosome 5, markers *D5S1456*, *D5S817* and *D5S488*. The markers on distal chromosome 5q were the only markers having $Z>3.0$, the criteria established for statistical significance. LOD scores greater than 2.0 ($\theta=0-0.15$) were also detected for chromosome 12 (*D12S391* $Z=2.03$, all families), and chromosome 21 (*D21S1440* $Z=2.62$, all families). The largest two-point LOD ($Z=4.3$; $\theta=0$) was observed for
25 marker *D5S408*, localized to chromosome 5q34-q35.

This initial chromosomal mapping was further refined by genotyping the three affected families for eight additional markers localized to region 5q34-q35. Six of these were informative for linkage (*D5S653*, *D5S498*, *D5S408*, *D5S2006*, *D5S1353* and *D5S1354*). Linkage analysis of these markers using VITESSE yielded a 2-point
30 LOD score of 6.1 at $\theta=0$ for marker *D5S1354* (Table 1) and a maximum multipoint LOD score of 8.8 at marker *D5S1354* (Fig. 2). These findings supported the localization of a gene within chromosome band 5q34-q35 that is a predisposing factor in hereditary lymphedema.

TABLE 1

**LOD scores for individual families estimated over
the interval defined by markers *D5S498* and *D5S2006*.**

		Z(θ) 0.0	Z(θ) 0.01	Z(θ)0.05	Z(θ) 0.1	Z(θ) 0.2
	<i>Locus D5S498</i>					
5	Family 101	-3.18	-2.33	-0.45	0.42	0.88
	Family 106	1.08	1.07	1.05	0.99	0.81
	Family 111	-0.85	-0.77	-0.53	-0.34	-0.13
	Family 105	1.22	1.20	1.11	0.98	0.72
	Family 135	-2.48	-1.85	-1.12	-0.75	-0.38
10	<i>Locus D5S1353</i>					
	Family 101	-2.99	-2.48	-1.21	-0.63	-0.18
	Family 106	0.28	0.29	0.35	0.38	0.38
	Family 111	-1.06	-1.02	-0.88	-0.72	-0.42
	Family 105	0.72	0.71	0.65	0.56	0.39
15	Family 135	-8.03	-4.18	-2.09	-1.13	-0.30
	<i>Locus D5S1354</i>					
	Family 101	6.09	6.02	5.69	5.21	4.07
	Family 106	1.42	1.40	1.32	1.20	0.96
	Family 111	0.21	0.22	0.23	0.24	0.22
20	Family 105	0.43	0.42	0.40	0.36	0.28
	Family 135	-6.88	-4.91	-3.20	-2.16	-1.07
	<i>Locus D5S408</i>					
	Family 101	2.80	2.74	2.50	2.20	1.56
	Family 106	0.66	0.68	0.73	0.76	0.71
25	Family 111	-1.70	-1.40	-0.80	-0.44	-0.10
	Family 105	0.42	0.41	0.38	0.35	0.27
	Family 135	-5.22	-4.24	-2.58	-1.67	-0.80

	$Z(\theta) 0.0$	$Z(\theta) 0.01$	$Z(\theta)0.05$	$Z(\theta) 0.1$	$Z(\theta) 0.2$
<i>Locus D5S2006</i>					
Family 101	4.51	4.70	4.85	4.66	3.80
Family 106	1.17	1.16	1.11	1.03	0.83
Family 111	-1.32	-1.18	-0.82	-0.56	-0.25
Family 105	0.43	0.42	0.40	0.36	0.28
Family 135	-3.86	-3.20	-2.11	-1.45	-0.73

During the completion of the genome scan, an additional ten lymphedema families were ascertained. Two of these families (Families 105 and 135, see Figures 1E and 1D), were potentially informative for linkage and were genotyped for markers in the linked region. Examination of the two point LOD scores for the five informative families for markers in the linked region (Table 1) shows that four of the families (101, 105, 106 and 111) are consistent with linkage to chromosome 5q while family 135 excluded linkage across the entire region with LOD scores $Z \leq -2.0$ for all markers. Multipoint linkage analysis of Families 101, 105, 106 and 111 (Fig. 2) yielded a peak LOD score of $Z = 10$ at marker *D5S1354*. These findings support the existence of at least two loci which predispose to hereditary lymphedema.

The order of markers *D5S1353*, *D5S1354* and *D5S408* with respect to each other was uncertain. Multipoint linkage analysis using alternative orders for these markers gave similar results. Marker *D5S498* is a framework marker and marker *D5S408* is mapped 11.2 centimorgans distal to *D5S498*, based on the CHLC chromosome 5 sex averaged, recombination minimized map, version 3 (www.chlc.org). The physical distance between *D5S498* and *D5S408* is estimated as 1.45 megabases based on the Genetic Location Database (LDB) chromosome 5 summary map (cedar.genetics.soton.ac.uk/public_html/).

Database analysis identified sixteen genes within this region. Two of these genes have been identified as having roles in development (*MSX2* and *VEGFR-3*). *MSX2* was considered an unlikely candidate gene for lymphedema because of its known involvement in craniofacial development [Jabs *et al.*, *Cell*, 75: 443-450 (1993)]. *VEGFR-3*, the gene encoding a receptor for VEGF-C, was selected as a better candidate gene for initial further study for the following reasons.

- (1) VEGFR-3 is expressed in developing lymphatic endothelium in the mouse [Kukk *et al.*, *Development*, 122: 3829-3837 (1996); and Kaipainen *et al.*, *Proc. Nat. Acad. Sci. USA*, 92: 3566-3570 (1995)];
- (2) expression of VEGFR-3 is induced in differentiating avian chorioallantoic membrane [Oh *et al.*, *Dev. Biol.*, 188:96-109 (1997)]; and
- (3) overexpression of VEGF-C, a ligand of VEGFR-3, leads to hyperplasia of the lymphatic vessels in transgenic mice [Jeltsch *et al.*, *Science*, 276: 1423-1425 (1997)].

10 To explore the potential role of VEGFR-3 in lymphedema, probands from the thirteen lymphedema families were screened for variation by direct sequencing of portions of the VEGFR-3 gene. The sequencing strategy used amplification primers generated based upon the *VEGFR-3* cDNA sequence (SEQ ID NO: 1) and information on the genomic organization of the related vascular endothelial growth factor receptor-2 (*VEGFR-2/KDR/flk-1*) [Yin *et al.*, *Mammalian Genome*, 9: 408-410 (1998)]. Variable positions (single nucleotide polymorphisms), the unique sequence primers used to amplify sequences flanking each variable site, and the method of detecting each variant are summarized in Table 2.

TABLE 2

Location, amplification primer sequences, amplification conditions, and detection methods for five intragenic single nucleotide polymorphisms in the human VEGFR-3 gene

Position in VEGFR-3 gene	Primer 1 sequence	Primer 2 sequence	Ann. temp.	[MgCl ₂]	Base change	Detection Method
Exon 12, amino acid 641	tcaccatcgatccaagc (SEQ ID NO: 7)	agttctgcgtgagccgag (SEQ ID NO: 8)	56 °C	1.0 mM	C→T	Sequencing
Exon 24, amino acid 1114	caggacgggggtgactga (SEQ ID NO: 9)	gccagggcctgtctactg (SEQ ID NO: 10)	56 °C	1.0 mM	C→T	Sequencing
Exon 3, amino acid 175	ccagctcctacgtgttcg (SEQ ID NO: 11)	ggcaacagctggatgtca (SEQ ID NO: 12)	56 °C	1.0 mM	C→T	<i>HhaI</i>
65bp 3' to Exon 6	ctgtgagggcggtgggagt (SEQ ID NO: 13)	gtcctttgagccactgga (SEQ ID NO: 14)	54 °C	1.5 mM	G→A	<i>SpyI</i>
55bp 3' to Exon 2	cacacgtcatcgacaccggtg (SEQ ID NO: 15)	ggcaacagctggatgtca (SEQ ID NO: 16)	56 °C	1.5 mM	C→T	<i>ApaI</i>

All amplifications were done for 35 cycles with denaturation at 94° for 30 seconds, annealing as above for 30 seconds, and extension at 72° for 30 seconds.

Amplification and sequencing primers were synthesized by the DNA Synthesis Facility, University of Pittsburgh. Amplification primers were tagged at the 5' end with the forward or reverse M13 universal sequence to facilitate direct sequencing. Amplimers were subjected to cycle sequencing using the dRhodamine terminator ready reaction kit or the Dye Primer ready reaction kit for -M13 and M13 Rev primers (Perkin Elmer) and analyzed on the Prism ABI 377 fluorescent sequencer. Sequences were aligned for further analysis using SEQUENCHER 3.0 (Gene Codes).

Genomic sequence from approximately 50% of the VEGFR-3 gene was determined in this manner, and five single nucleotide variants were observed. Two of the variants occurred in introns, and a third was a silent substitution in predicted exon 3. These intragenic polymorphisms were used to map the *VEGFR-3* gene. As shown in Figure 2, *VEGFR-3* maps within the region of chromosome 5q linked to the lymphedema phenotype, consistent with it being selected as a candidate gene. In two families, (Family 127, pedigree not shown, and Family 135), a C→T transition was identified at nucleotide position 1940 of the *VEGFR-3* cDNA (SEQ ID NO: 1). This nucleotide substitution is predicted to lead to a non-conservative substitution of serine (codon TCC) for proline (codon CCC) at residue 641 (putative exon 12, within the sixth immunoglobulin-like region of the receptor's extracellular domain) of the amino

acid sequence of the receptor (SEQ ID NO: 2). However, this sequence change was observed in 2 of 120 randomly selected individuals from the general population (240 alleles). Also, in one of the two families in which this variant was initially detected, family 135, linkage between lymphedema and chromosome 5q markers was excluded (Table 1 and Figure 2). In probands from the other ten families, wild type sequence was observed at nucleotide position 1940. Collectively, these results suggest that this P641S variant is not causative.

In one nuclear family (Family 127, pedigree shown in Figure 1F) a C-T transition was observed at nucleotide position 3360 (SEQ ID NO: 1) of the *VEGFR-3* cDNA. This nucleotide substitution is predicted to lead to a non-conservative substitution of leucine (codon CTG) for proline (codon CCG) at residue 1114 of the amino acid sequence of the receptor (SEQ ID NO: 2). This P1114L mutation is predicted to lie in the intracellular tyrosine kinase domain II involved in intracellular signaling [Pajusola *et al.*, *Cancer Res.*, 52:5738- 5743 (1992)]. Direct sequencing of predicted exon 24 of the *VEGFR-3* gene alleles from members of this family identified this substitution only in affected and at-risk family members. This sequence change was not observed in 120 randomly selected individuals of mixed European ancestry from the general population (240 alleles). In probands from the other 11 families, wild type sequence was observed at nucleotide position 3360.

Collectively, this data demonstrates that a missense mutation that causes a non-conservative substitution in a kinase domain of the VEGFR-3 protein correlates strongly with a heritable lymphedema in one family, and suggests that other mutations in the same gene may exist that correlate with heritable lymphedema in other families. As explained above, only a portion of the VEGFR-3 gene sequence was analyzed to identify this first mutant of interest. Additional sequencing, using standard techniques and using the known *VEGFR-3* gene sequence for guidance, is expected to identify additional mutations of interest that are observed in affected and at-risk members of other families studied.

EXAMPLE 2

Demonstration that a C→T missense mutation at position 3360 in the VEGFR-3 coding sequence results in a tyrosine kinase negative mutant

The results set forth in Example 1 identified two missense mutations in
5 the *VEGFR-3* coding sequence, one of which (C→T at position 3360) appeared to
correlate with heritable lymphedema and one of which (C→T transition at position
1940) did not. The following experiments were conducted to determine the
biochemical significance of these mutations on VEGFR-3 biological activity.

To analyze how the two single amino acid substitutions affect the
10 VEGFR-3-mediated signaling, the corresponding mutant receptor expression vectors
were generated using site-directed mutagenesis procedures and expressed in 293T cells
by transient transfection. The long form of human VEGFR-3 cDNA (SEQ ID NO: 1)
was cloned as a *Hind* III-*Bam* HI fragment from the LTR-FLT4l plasmid [Pajusola *et*
al., *Oncogene* 8: 2931-2937 (1993)] into pcDNA3.1/Z(+) (Invitrogen). The P641S
15 and P1114L mutants of VEGFR-3 were generated from this construct with the
GeneEditor™ *in vitro* Site-Directed Mutagenesis System (Promega) using the
following oligonucleotides (the C→T mutations are indicated with bold letters):

5'-CCTGAGTATC**T**CCCGCGTCGC-3' (SEQ ID NO: 17) for P641S
mutation; and
20 5'-GGTGCCTCC**C**TGTACCCTGGG-3' (SEQ ID NO: 18) for P1114L
mutation.

For the transient expression studies, 293T cells were grown in
Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf
serum (GIBCO BRL, Life Technologies, Gaithersburg, MD), glutamine, and
25 antibiotics. Cells were transfected with 20 µg of plasmid encoding the wild type or
mutant VEGFR-3 forms using the calcium phosphate method, and harvested 36 hours
after transfection for immunoprecipitation and Western blotting. Under these
conditions, RTK overexpression results in ligand-independent activation, thus allowing
the receptor phosphorylation to be studied. An empty vector was used for mock
30 (control) transfections. (It will be appreciated that ligand stimulation assays of
VEGFR-3 forms also can be employed, *e.g.*, as described in U.S. Patent No.
5,776,755, incorporated herein by reference, using VEGF-C or VEGF-D ligands.)

In order to investigate the effect of the two VEGFR-3 mutants on the tyrosine phosphorylation of the VEGFR-3, Western blotting analysis was performed using anti-phosphotyrosine antibodies. The cell monolayers were washed three times with cold phosphate-buffered saline (PBS, containing 2 mM vanadate and 2 mM PMSF) and scraped into RIPA buffer (150 mM NaCl, 1 % Nonidet P40, 0.5 % deoxycholic acid sodium salt, 0.1 % SDS, 50 mM Tris-HCl, pH 8.0) containing 2 mM Vanadate, 2 mM PMSF, and 0.07 U/ml Aprotinin.

The cell lysates were sonicated and centrifuged for 10 minutes at 19,000 X g, and the supernatants were incubated for 2 hours on ice with 2 µg/ml of monoclonal anti-VEGFR-3 antibodies (9D9f9) [Jussila *et al.*, *Cancer Res.*, 58: 1599-604 (1998)]. Thereafter, Protein A sepharose (Pharmacia) beads were added and incubation was continued for 45 minutes with rotation at +4°C. The sepharose beads were then washed three times with ice-cold RIPA buffer and twice with PBS (both containing 2 mM vanadate, 2 mM PMSF), analyzed by 7.5 % SDS-PAGE and transferred to a nitrocellulose filter (Protran Nitrocellulose, Schleicher & Schuell, No. 401196) using semi-dry transfer apparatus. After blocking the filter with 5 % BSA in TBS-T buffer (10 mM Tris, pH 7.5, 150 mM NaCl, 0.05 % Tween 20), the filters were incubated with the phosphotyrosine-specific primary antibodies (Upstate Biotechnology, #05-321), followed by biotinylated goat-anti-mouse immunoglobulins (Dako, E0433) and Biotin-Streptavidin HRP complex (Amersham, RPN1051). The bands were visualized by the enhanced chemiluminescence (ECL) method.

After analysis for phosphotyrosine-containing proteins, the filters were stripped by washing for 30 minutes at +50°C in 100 mM 2-mercaptoethanol, 2 % SDS, 62.5 mM Tris-HCl, pH 6.7, with occasional agitation. The filters were washed with TBS-T, blocked again with BSA as described above, and analyzed for the presence of VEGFR-3 using the 9D9f9 antibodies and HRP-conjugated rabbit-anti-mouse immunoglobulins (Dako, P0161).

The Western analyses revealed that the P641S mutant receptor was phosphorylated normally, *i.e.*, in a manner similar to the wild type control. However, the proteolytic processing of the P641S receptor protein may be affected, as the 175 kD and 125 kD polypeptides seemed to have a higher relative density when compared to the 195 kD form.

In contrast, no phosphorylated P1114L mutant protein was detected using the phosphotyrosine antibodies. The expression of similar amounts of the VEGFR-3 protein (normal and both mutants) was confirmed using the monoclonal 9D9f9 antibody, which is directed towards the extracellular domain of the VEGFR-3.

5 Both the P641S and the P1114L mutant VEGFR-3 migrated slightly faster than the wild type VEGFR-3 in the gel electrophoresis.

In order to analyze the possible dominant negative effect of the P1114L mutant on the wild-type receptor, a second, similar set of experiments were performed wherein the 293T cells were transfected with an increasing amount of the P1114L expression vector in combination with decreasing amounts of the wild type vector. Wild type to mutant ratios of 1:0, 3:1, 1:1, 1:3 and 0:1 were used. The cells were lysed 48 hours after transfection and the lysates were analyzed by immunoprecipitation and Western blotting as described above. These experiments permitted evaluation of whether the mutant protein interferes with wild type protein phosphorylation and estimation of the minimal amount of the WT protein needed for observable tyrosyl autophosphorylation. Immunoprecipitates from cells transfected with only the WT plasmid revealed WT protein that was strongly phosphorylated in this experiment (lane 2), whereas immunoprecipitates from cells transfected with only the mutant plasmid were again inactive (unphosphorylated).

10

15

Interestingly, when transfection was made using 75% of WT and 25 % of mutant plasmid, the phosphorylation of the receptors was decreased by about 90%. This result strongly suggests that the P1114L mutant receptor forms heterodimers with the WT receptor, but cannot phosphorylate the WT receptor, thus failing to activate it. Under this theory, the WT receptor monomers in the heterodimers would also remain inactive, causing a disproportionate decrease of the total amount of activated receptor, when co-transfected with the mutant. Wildtype-wildtype homodimers would remain active and be responsible for the observed signaling. When the wild type and mutant receptor expression vectors were transfected at a 1:1 ratio, the VEGFR-3 phosphorylation was about 4% of the wild type alone, whereas at a 1:3 ratio, no tyrosine phosphorylation of VEGFR-3 was observed.

20

25

30

The foregoing results are consistent with the linkage analyses in Example 1: the mutation at position 641 that did not appear to correlate with lymphedema also did not appear to be dysfunctional, whereas the mutation at position

1114 appeared to cause a dominant negative mutation that shows no tyrosine phosphorylation alone and that drastically reduces VEGFR-3 signaling in cells expressing both the mutant and wild type VEGFR-3 genes.

Collectively, these data indicate that the P1114L VEGFR-3 mutant is unable to act as a part of the signaling cascade, and also acts in a dominant negative manner, thus possibly interfering partially with the activation of the wild type VEGFR-3. Such effects of the mutation may eventually lead to lymphedema.

EXAMPLE 3

Treatment of lymphedema with a VEGFR-3 ligand

The data from Examples 1 and 2 collectively indicate a causative role in heritable lymphedema for a mutation in the *VEGFR-3* gene that interferes with VEGFR-3 signaling. Such a mutation behaves in an autosomal dominant pattern, due to the apparent necessity for receptor dimerization in the signaling process. However, the data from Example 2 suggests that some residual signaling may still occur in heterozygous affected individuals, presumably through pairing of VEGFR-3 proteins expressed from the wild type allele. The following experiments are designed to demonstrate the efficacy of VEGFR-3 ligand treatment in such affected individuals, to raise VEGFR-3 signaling to levels approaching normal and thereby ameliorate/palliate the symptoms of hereditary lymphedema.

Initially, an appropriate animal model is selected. Several potential animal models have been described in the literature. [See, e.g., Lyon *et al.*, *Mouse News Lett.* 71: 26 (1984), *Mouse News Lett.* 74: 96 (1986), and *Genetic variants and strains of the laboratory mouse*, 2nd ed., New York: Oxford University Press (1989), p. 70 (*Chylous ascites* mouse); Dumont *et al.*, *Science*, 282: 946-949 (1998) (heterozygous VEGFR-3 knockout mouse); Patterson *et al.*, "Hereditary Lymphedema," *Comparative Pathology Bulletin*, 3: 2 (1971) (canine hereditary lymphedema model); van der Putte, "Congenital Hereditary Lymphedema in the Pig," *Lympho*, 11: 1-9 (1978); and Campbell-Beggs *et al.*, "Chyloabdomen in a neonatal foal," *Veterinary Record*, 137: 96-98 (1995).] Those models which are determined to have analogous mutations to the VEGFR-3 gene are preferred. Analogous mutations would include mutations affecting corresponding residues and also mutations affecting different residues but causing similar functional alterations. The *Chylous ascites*

mouse VEGFR-3 gene contains a missense mutation at a position corresponding to residue 1053 of SEQ ID No. 2, which maps to the catalytic pocket region of the tyrosine kinase catalytic domain. Thus, the “*Chy*” mouse is expected to display similar functional alterations to human mutations affecting tyrosine kinase activity, a prediction which can be confirmed by functional assays such as those described in Example 2. In a preferred embodiment, “knock in” homologous recombination genetic engineering strategies are used to create an animal model (*e.g.*, a mouse model) having a VEGFR-3 allelic variation analogous to the human variations described herein. [See, *e.g.*, Partanen *et al.*, *Genes & Development*, 12: 2332-2344 (1998) (gene targeting to introduce mutations into another receptor protein (FGFR-1) in mice).] For example, the P1114L mutation in human VEGFR-3 occurs in a VEGFR-3 region having highly conserved amino acid identity with murine VEGFR-3 (Genbank Accession No. L07296). Thus, a corresponding P1114L can be introduced into the murine VEGFR-3 by “knock-in” homologous recombination. Optionally, such mice can be bred to the heterozygous VEGFR-3 knockout mice or *Chy* mice described above to further modify the phenotypic severity of the lymphedema disease.

The mice as described above are treated with a candidate therapeutic, *e.g.*, a recombinant mature form of VEGF-C, at various dosing schedules, *e.g.*, once daily by intravenous (IV) or intramuscular (IM) injection at a dose of 1-1000 ng/g body weight, preferably 10-100 ng/g, which should result in a peak level saturating VEGFR-3 (K_d about 150 pM) but not VEGFR-2 (K_d around 400 pM). For VEGFR-3-specific forms, such as VEGF-C Δ C₁₅₆, even higher dosing is contemplated, to sustain VEGFR-3-saturating physiological concentrations for longer periods. Direct IM injection at multiple sites in the muscles of affected extremities is a preferred route of administration. The dosing is adjusted according to the efficacy of the treatment and the presence of possible side effects due to the lowering of blood pressure, which has been observed in response to VEGF administration IV. The efficacy of treatment is measured via NMRI imaging of the water content and volume of swelling of the abdomen and the extremities of the animals. The amount of fluid in the abdominal cavity is estimated and the animals are weighed during the follow-up.

In studies using VEGFR-3 $-/+ \times Chy$ mice progeny, the animals will also have the β -galactosidase marker in their lymphatic endothelium. After a successful treatment, the treated and non-treated experimental animals and VEGFR-3

-/+ controls are killed and their lymphatic vessels are visualized by β -gal and antibody staining. The staining patterns of experimental and control animals are compared for vessel diameter, numbers of endothelial cells, density of blood and lymphatic vessels, and nuclear density/section surface area for the estimation of tissue oedema.

- 5 Such experiments are repeated with various candidate therapeutics (e.g., VEGF-C or VEGF-D recombinant polypeptides; VEGF-C and VEGF-D gene therapy vectors; and combinations thereof) at various dosing schedules to determine an optimum treatment regimen.

EXAMPLE 4

Chromosomal structure of the human *VEGFR-3* gene

Sequencing and mapping of human DNA corresponding to the *VEGFR-3* locus has indicated that this gene consists of thirty exons separated by twenty-nine introns of varying size. The exon intron organization is summarized as follows:

	<u>EXON NUMBER</u>	<u>Bp of SEQ ID NO: 1 size (bp)</u>	<u>INTRON SIZE</u>
15	1	20-77 58 bp	unknown
	2	78-174 97 bp	>1kb
20	3	175-419 245 bp	218 bp
	4	420-532 113 bp	120 bp
25	5	533-695 163 bp	107 bp
	6	696-835 140 bp	269 bp
	7	836-1004 169 bp	261 bp
30	8	1005-1122 118 bp	>1 kb
	9	1123-1277	

		155 bp	unknown
	10	1278-1440 163 bp	>1 kb
5	11	1441-1567 127 bp	unknown
	12	1568-1676 109 bp	unknown
	13	1677-2039 363 bp	293 bp
10	14	2040-2186 147 bp	99 bp
	15	2187-2318 132 bp	approx. 160 bp
15	16	2319-2425 107 bp	301 bp
	17	2426-2561 139 bp	>464 bp
	18	2562-2666 105 bp	unknown
20	19	2667-2780 114 bp	143 bp
	20	2781-2869 89 bp	>1 kb
25	21	2870-3020 151 bp	unknown
	22	3021-3115 95 bp	unknown
	23	3116-3238 123 bp	unknown
30	24	3239-3350 112 bp	974 bp
	25	3351-3450 100 bp	400 bp

	26	3451-3557 107 bp	unknown
	27	3558-3705 148 bp	>1 kb
5	28	3706-3826 121 bp	unknown
	29	3827-3912 86 bp	unknown
10	30a (Flt4 short)	3913-4111 199 bp	3.7 kb
	30b (Flt4long)	3913-4416 >504 bp	(CDS 504 bp)

15 The foregoing information permits rapid design of oligonucleotides for amplifying select portions of the VEGFR-3 gene from genomic DNA, or RNA, or cDNA, to facilitate rapid analysis of an individual's *VEGFR-3* coding sequence, to determine whether the individual possesses a mutation that correlates with a lymphedema phenotype.

EXAMPLE 5

Identification of additional non-conservative missense mutants

20 Using procedures essentially as described in Example 1, the VEGFR-3 coding sequences from additional affected and unaffected individuals from families having members suffering from heritable lymphedema were studied. The analysis focused on families with statistical linkage to chromosome 5q as described in Example 1. The additional analysis included the PCR amplification and sequencing of Exon 17, 25 Exon 22, and Exon 23 sequences with the following PCR primers:

	Exon 17-1	5'-CATCAAGACGGGCTACCT-3' (SEQ ID NO: 23)
	Exon 17-2	5'-CCGCTGACCCACACCTT-3' (SEQ ID NO: 24)
	Exon 22-1	5'-GAGTTGACCTCCCAAGGT-3' (SEQ ID NO: 25)
	Exon 22-2	5'-TCTCCTGGACAGGCAGTC-3' (SEQ ID NO: 26)
30	Exon 23-1	5'-GAGTTGACCTCCCAAGGT-3' (SEQ ID NO: 27)
	Exon 23-2	5'-TCTCCTGGACAGGCAGTC-3' (SEQ ID NO: 28)

These additional studies identified four additional non-conservative missense mutations in evolutionarily conserved amino acids in kinase domains I and II of human VEGFR-3. Each mutation, shown in Table 3 below, was observed in a single independently ascertained family, and in each family, the mutation co-segregates with individuals suffering from, or considered at risk for developing, lymphedema. None of these mutations were observed in the *VEGFR-3* genes in a random sample of more than 300 chromosomes from individuals from families unaffected with heritable lymphedema.

TABLE 3
Mutations in VEGFR-3 causing Hereditary Lymphedema*

Exon	Nucleotide Substitution**	Amino Acid Substitution	Functional Domain
24	C3360T	P1114L	Kinase 2
17	G2588A	G857R	Kinase 1
23	G3141C	R1041P	Kinase 2
23	T3150C	L1044P	Kinase 2
23	G3164A	D1049N	Kinase 2

* Numbers indicate nucleotide or amino acid positions in SEQ ID NOs: 1 and 2.

**It will be appreciated that, since DNA is double-stranded, each mutation could be characterized in two equivalent ways, depending on whether reference is being made to the coding or the non-coding strand.

Referring to SEQ ID NO: 2, the kinase domains of VEGFR-3 comprise approximately residues 843-943 and residues 1009-1165. Within these domains, molecular modeling suggests that residues G852, G854, G857, K879, E896, H1035, D1037, N1042, D1055, F1056, G1057, E1084, D1096 and R1159 are of particular importance in comprising or shaping the catalytic pocket within the kinase domains. See van Der Geer and Hunter, *Ann. Rev. Cell. Biol.*, 10: 251-337 (1994); and Mohammadi *et al.*, *Cell* 86: 577-587 (1996). Thus, this data identifying additional mutations implicate missense mutations within a kinase domain of the VEGFR-3 protein as correlating strongly with a risk for developing a heritable lymphedema phenotype. Mutations which affect residues in and around the catalytic pocket appear particularly likely to correlate with lymphedema. The P1114L mutation, though not

situated within the catalytic pocket, is postulated to cause a conformational alteration that affects the catalytic pocket. The G857R mutation is postulated to block the catalytic pocket and/or the ATP binding site of the kinase domain.

EXAMPLE 6

5 Functional Analysis of Additional *VEGFR-3* missense mutations

Using procedures essentially as described above in Example 2, the functional state of the G857R, L1044P, and D1049N mutations were analyzed. (PLCLB buffer, comprising 150 mm NaCl, 5% glycerol, 1% Triton X-100, 1.5M MgCl₂, 50 mm HEPES, pH 7.5, was substituted for RIPA buffer described in Example 2 for immunoprecipitation and Western blotting protocols.) A *VEGFR-3*-encoding construct comprising the G857R mutation was generated from the long form of human *VEGFR-3* cDNA using the oligonucleotide:

5'-CGG CGC CTT CAG GAA GGT GGT-3' (SEQ ID NO: 20)

A construct comprising the L1044P mutation was generated from the long form of human *VEGFR-3* cDNA using the oligonucleotide:

5'-CGG AAC ATT CCG CTG TCG GAA-3' (SEQ ID NO: 21)

A construct comprising the D1049N mutation was generated from the long form of human *VEGFR-3* cDNA using the oligonucleotide:

5'-GTC GGA AAG CAA CGT GGT GAA-3' (SEQ ID NO: 22).

The constructs were transiently transfected into 293T cells and harvested for Western blotting essentially as described in Example 2, except for the buffer substitution described above. In contrast to wild type *VEGFR-3* and *VEGFR-3* containing the P641S mutation, no phosphorylated G857R or L1044P mutant protein was detected using the phosphotyrosine antibodies, consistent with the results that had been observed for P1114L. The expression of similar amounts of the *VEGFR-3* protein was confirmed using the monoclonal 9D9f9 antibody, which is directed towards the extracellular domain of the *VEGFR-3* in the Western blotting. This data suggested that these observed mutations did indeed affect *VEGFR-3* kinase function. The D1049N mutant appeared to retain at least some tyrosine kinase activity. It is also noteworthy that *VEGFR-1* and *VEGFR-2* contain an asparagine residue at the position in their tyrosine kinase domains which corresponds to position 1049 of *VEGFR-3*.

Together, these data suggest that the D1049N variation may only be an allelic variant that correlates with hereditary lymphedema, rather than a causative mutation.

To determine whether the VEGFR-3 mutants function in a dominant negative manner, each construct was co-transfected at varying ratios with wild type receptor into 293T cells essentially as described in Example 2. Unlike the results observed for P1114L and described in Example 2, neither the G857R mutant nor the L1044P mutant seemed to interfere with phosphorylation of the co-transfected wild type receptor.

The absence of a dominant negative effect in these experiments does not foreclose a conclusion that the mutations described above are causative. It has been found that a significant fraction of ligand-activated receptor tyrosine kinases traffic to the lysosomal compartment after internalization, where they are degraded. However, receptors which are not ligand-activated preferentially recycle back to the cell surface after internalization. Thus, it is possible that the turnover time of the weakly phosphorylated mutant receptor is significantly longer than that of the wild type receptor protein. If this were true, the amount of the mutant receptor on the endothelial cell surface could be considerably higher than the amount of the phosphorylated and rapidly internalized wild type receptor, and any available ligand would thus bind a disproportionately high number of mutant receptors. Both a possible dominant negative effect of the mutant receptor and an abnormally long half-life of the tyrosine kinase negative mutant receptor could eventually lead to lymphedema. Alternatively, a mutation that merely decreases (but does not eliminate) VEGFR-3 tyrosine kinase activity may display a constitutive low level of internalization and degradation that is insufficient to trigger sufficient downstream signalling, but decreases the effective concentration of VEGFR-3 on cell surfaces for ligand binding and effective activation, leading eventually to lymphedema.

While the present invention has been described in terms of specific embodiments, it is understood that variations and modifications will occur to those in the art, all of which are intended as aspects of the present invention. Accordingly, only such limitations as appear in the claims should be placed on the invention.

CLAIMS

What is claimed is:

1. A method of screening a human subject for an increased risk of developing a lymphatic disorder, comprising the steps of:

5 (a) assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering the sequence or expression of at least one VEGFR-3 allele; and

(b) screening for an increased risk of developing a lymphatic disorder from the presence or absence of said mutation, wherein the presence of a mutation altering the encoded amino acid sequence or expression of at least one VEGFR-3 allele in the
10 nucleic acid correlates with an increased risk of developing a lymphatic disorder.

2. A method according to claim 1 wherein the assaying step comprises determining the presence or absence of a mutation altering a tyrosine kinase domain amino acid sequence of the protein encoded by the VEGFR-3 allele.

15 3. A method according to claim 1 wherein the assaying step comprises determining the presence or absence of a missense mutation in a VEGFR-3 allele at a position corresponding to one of codons 857, 1041, 1044 and 1049 of the VEGFR-3-encoding sequence set forth in SEQ ID NO: 1.

20 4. A method according to claim 1 wherein the assaying step comprises determining the presence or absence of a missense mutation in the VEGFR-3 allele at a position corresponding to codon 1114 of the VEGFR-3-encoding sequence set forth in SEQ ID NO: 1.

5. A method according to claim 1 wherein the assaying step comprises at least one procedure selected from the group consisting of:

(a) determining a nucleotide sequence of at least one codon of at least one VEGFR-3 allele of the human subject;

5 (b) performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences;

(c) performing a polynucleotide migration assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different
10 from one or more reference sequences; and

(d) performing a restriction endonuclease digestion to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences.

6. A method according to claim 1 wherein the assaying step comprises:
15 performing a polymerase chain reaction (PCR) to amplify nucleic acid comprising VEGFR-3 coding sequence, and determining nucleotide sequence of the amplified nucleic acid.

7. A method of screening for a VEGFR-3 hereditary lymphedema genotype in a human patient, comprising the steps of:

20 (a) providing a biological sample comprising nucleic acid from said patient, said nucleic acid including sequences corresponding to said patient's VEGFR-3 alleles;

(b) analyzing said nucleic acid for the presence of a mutation or mutations;

25 (c) determining a VEGFR-3 genotype from said analyzing step; and

(d) correlating the presence of a mutation in a VEGFR-3 allele with a hereditary lymphedema genotype.

8. The method according to claim 7 wherein said biological sample is a cell sample.

9. The method according to claim 7 wherein said analyzing comprises sequencing a portion of said nucleic acid, said portion comprising at least one codon of said VEGFR-3 alleles.

10. The method according to claim 7 wherein said nucleic acid is DNA.

5 11. The method according to claim 7 wherein said nucleic acid is RNA.

12. A method of treatment for hereditary lymphedema, comprising the step of administering to a patient with hereditary lymphedema a therapeutically effective amount of a growth factor product selected from the group consisting of vascular endothelial growth factor C (VEGF-C) protein products, vascular endothelial growth factor D (VEGF-D) protein products, VEGF-C gene therapy products, and VEGF-D gene therapy protein products.

13. A therapeutic or prophylactic method of treating lymphedema, comprising the steps of:
providing isolated lymphatic endothelial cells or lymphatic endothelial progenitor cells;
transforming or transfecting the cells *ex vivo* with a polynucleotide comprising a nucleotide sequence that encodes a wild type VEGFR-3;
and administering the transformed or transfected cells to the mammalian subject.

14. An oligonucleotide useful as a probe for identifying polymorphisms in a human Flt4 receptor tyrosine kinase gene, the oligonucleotide comprising 6-50 nucleotides that have a sequence that is identical or exactly complementary to a portion of a wild type human VEGFR-3 gene sequence or VEGFR-3 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution.

15. An oligonucleotide according to claim 14 wherein the nucleotide sequence is exactly identical or exactly complementary to a portion of the human VEGFR-3 coding sequence set forth in SEQ ID NO: 1, except for a nucleotide substitution at a position corresponding to a nucleotide selected from the group consisting of bases 2546 through 2848 and bases 3044 through 3514.

16. An oligonucleotide according to claim 14 wherein the nucleotide sequence is exactly identical or exactly complementary to a portion of the human VEGFR-3 coding sequence set forth in SEQ ID NO: 1, except for a nucleotide substitution at a position corresponding to nucleotide 3360 of SEQ ID NO: 1.

17. An oligonucleotide according to claim 14 wherein the nucleotide sequence is exactly identical or exactly complementary to a portion of the human VEGFR-3 coding sequence set forth in SEQ ID NO: 1, except for a nucleotide substitution at a position corresponding to a nucleotide selected from the group consisting of position 2588, position 3141, position 3150 and position 3164 of SEQ ID NO: 1.

18. A kit comprising at least two oligonucleotides of the formula X_nYZ_m or its complement;

where n and m are integers from 0 to 49;

where $5 \leq (n + m) \leq 49$;

where X_n is a stretch of n nucleotides identical to a first portion of SEQ ID NO: 1, said first portion ending immediately upstream (5') of position 3360 of SEQ ID NO: 1; and

where Z_m is a stretch of m nucleotides identical to a second portion of SEQ ID NO: 1, said second portion beginning immediately downstream (3') of position 3360 of SEQ ID NO: 1; and

wherein Y represents a nucleotide selected from the group consisting of adenine, guanine, cytosine, thymine, and uracil nucleotides.

19. A kit comprising at least two oligonucleotides of the formula X_nYZ_m or its complement;

where n and m are integers from 0 to 49;

where $5 \leq (n + m) \leq 49$;

5 where X_n is a stretch of n nucleotides identical to a first portion of SEQ ID NO: 1, said first portion ending immediately upstream (5') of position W of SEQ ID NO: 1; and

10 where Z_m is a stretch of m nucleotides identical to a second portion of SEQ ID NO: 1, said second portion beginning immediately downstream (3') of position W of SEQ ID NO: 1;

wherein position W of SEQ ID NO: 1 is selected from the group consisting of nucleotides 2588, 3141, 3150, and 3164 of SEQ ID NO: 1; and

wherein Y represents a nucleotide selected from the group consisting of adenine, guanine, cytosine, thymine, and uracil nucleotides.

15 20. An array of oligonucleotide probes immobilized on a solid support, wherein each probe occupies a separate known site in the array; and wherein the array includes at least one probe set comprising two to four probes, wherein one probe is exactly identical or exactly complementary to a wild type human VEGFR-3 coding sequence, and the other one to three members of the set are exactly identical to the
20 first member, but for at least one different nucleotide, which different nucleotide is located in the same position in each of the one to three additional set members.

21. An array of oligonucleotide probes immobilized on a solid support according to claim 20, wherein each probe occupies a separate known site in the array; and wherein the array includes at least one probe set comprising two to four probes,
25 wherein one probe is exactly identical or exactly complementary to a portion of a human VEGFR-3 coding sequence set forth in SEQ ID NO: 1, and the other one to three members of the set are exactly identical to the first member, but for at least one different nucleotide, which different nucleotide is located in the same position in each of the one to three additional set members, said position corresponding to a position
30 selected from the group consisting of bases 2546 through 2848 and bases 3044 through 3514 of SEQ ID NO: 1.

22. A purified polynucleotide comprising a nucleotide sequence encoding a human VEGFR-3 protein variant, wherein said polynucleotide is capable of hybridizing to the complement of SEQ ID NO: 1 under the following hybridization conditions:

hybridization at 42°C in 50% formamide, 5X SSC, 20 mM Na•PO₄, pH 6.8;

5 and

washing in 0.2X SSC at 55°C;

and wherein the encoded VEGFR-3 protein variant has an amino acid sequence that differs from the amino acid sequence set forth in SEQ ID NO: 2 at one or more positions selected from the group consisting of amino acids 843 to 943 of SEQ ID
10 NO: 2 and amino acids 1009 to 1165 of SEQ ID NO: 2.

23. A purified polynucleotide according to claim 22, wherein the encoded VEGFR-3 protein variant has an amino acid sequence that differs at position 1114 from the amino acid sequence set forth in SEQ ID NO: 2.

24. A purified polynucleotide according to claim 22 wherein the encoded
15 VEGFR-3 protein variant has an amino acid sequence that differs from the amino acid sequence set forth in SEQ ID NO: 2 at position selected from the group consisting of residues 857, 1041, 1044 and 1049 of SEQ ID NO: 2.

25. A purified polynucleotide comprising a nucleotide sequence encoding a VEGFR-3 protein of a human that is affected with heritable lymphedema;

20 wherein said polynucleotide is capable of hybridizing to the complement of SEQ ID NO: 1 under the following hybridization conditions: hybridization at 42°C in 50% formamide, 5X SSC, 20 mM Na•PO₄, pH 6.8; and washing in 0.2X SSC at 55°C;

25 and wherein the polynucleotide encodes a VEGFR-3 amino acid sequence that differs from SEQ ID NO: 2 at at least one residue.

26. A purified polynucleotide according to claim 25 wherein the polynucleotide encodes an amino acid sequence that differs from SEQ ID NO: 2 at at least one residue selected from the group consisting of residues 843 to 943 and 1009 to 1165 of SEQ ID NO: 2.

5 27. A vector comprising a polynucleotide according to claim 25.

28. A host cell that has been transformed or transfected with a polynucleotide according to claim 25 and that expresses the VEGFR-3 protein encoded by the polynucleotide.

10 29. A host cell according to claim 28 that has been co-transfected with a polynucleotide encoding the VEGFR-3 amino acid sequence set forth in SEQ ID NO: 2 and that expresses the VEGFR-3 protein having the amino acid sequence set forth in SEQ ID NO: 2.

30. A method for identifying a modulator of intracellular VEGFR-3 signaling, comprising the steps of:
15 a) contacting a cell expressing at least one mutant mammalian VEGFR-3 polypeptide in the presence and in the absence of a putative modulator compound;
b) detecting VEGFR-3 signaling in the cell; and
c) identifying a putative modulator compound in view of decreased
20 or increased signaling in the presence of the putative modulator, as compared to signaling in the absence of the putative modulator.

31. A method according to claim 30 wherein the cell expresses the mutant mammalian VEGFR-3 polypeptide and a wildtype mammalian VEGFR-3 polypeptide.

25 32. A method according to claim 31 wherein the mutant and wildtype VEGFR-3 polypeptides are human.

33. A method according to claim 32 wherein said mutant VEGFR-3 polypeptide is characterized by a substitution or deletion mutation in a kinase domain of the VEGFR-3 polypeptide.

5 34. A method according to claim 32 wherein said mutant VEGFR-3 polypeptide is characterized by at least one substitution or deletion of the wild type VEGFR-3 amino acid sequence set forth in SEQ ID NO: 2, said at least one substitution or deletion occurring at a position corresponding to a residue selected from positions 843 to 943 and positions 1009 to 1165 of SEQ ID NO: 2.

10 35. A method according to claim 32 wherein the mutant VEGFR-3 polypeptide comprises a leucine amino acid at the position corresponding to position 1114 of SEQ ID NO: 2.

15 36. A method according to claim 32 wherein said mutant VEGFR-3 polypeptide is characterized by at least one substitution or deletion of the wild type VEGFR-3 amino acid sequence set forth in SEQ ID NO: 2, said at least one substitution or deletion occurring at a position corresponding to a residue selected from positions 857, 1041, 1044 and 1049, and 1114 of SEQ ID NO: 2.

The present invention provides materials and methods for screening for and treating hereditary lymphedema in human subjects.

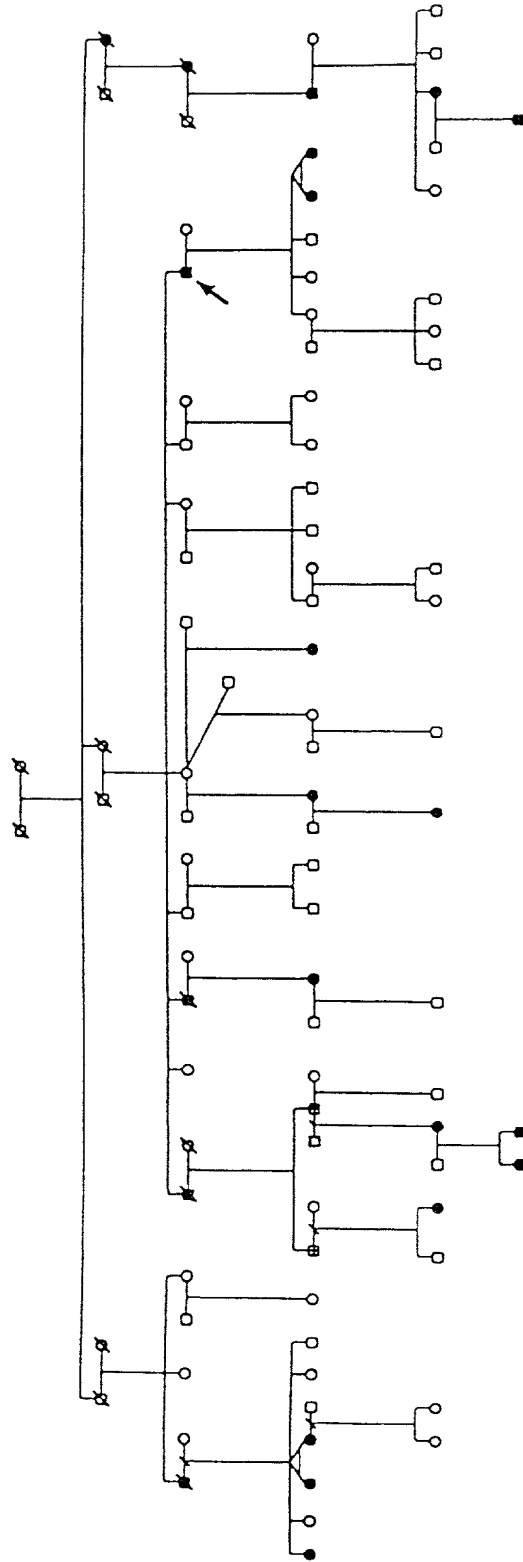


Fig. 1A

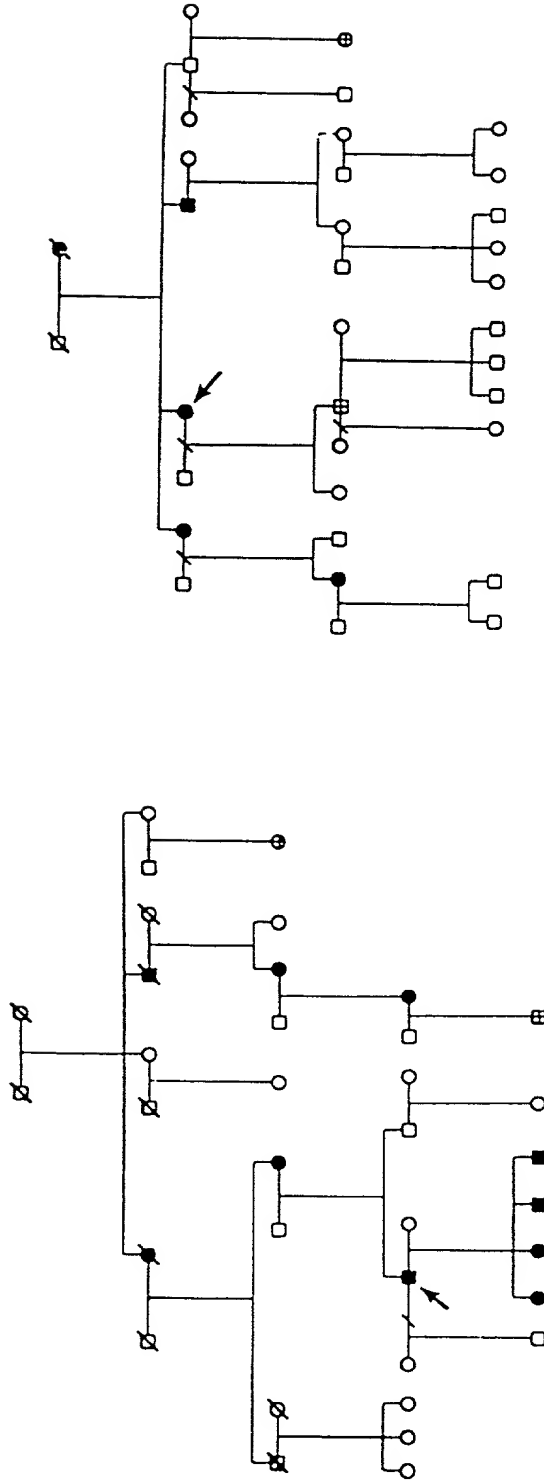


Fig. 1C

Fig. 1B

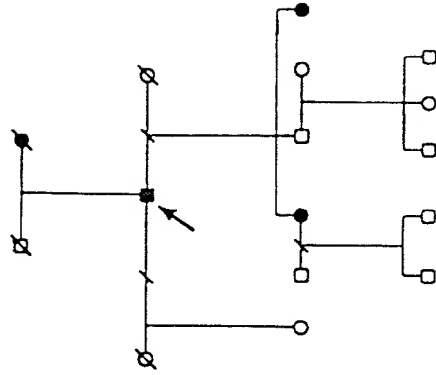


Fig. 1E

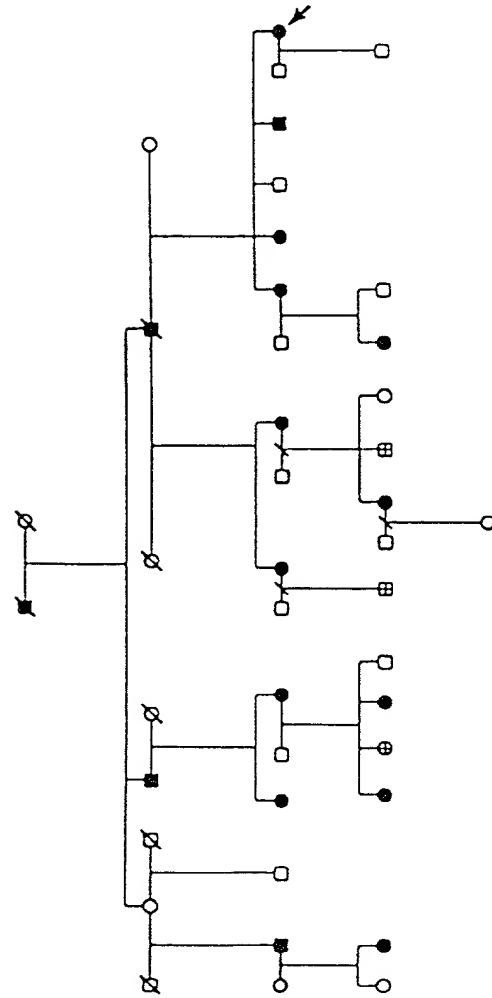


Fig. 1D

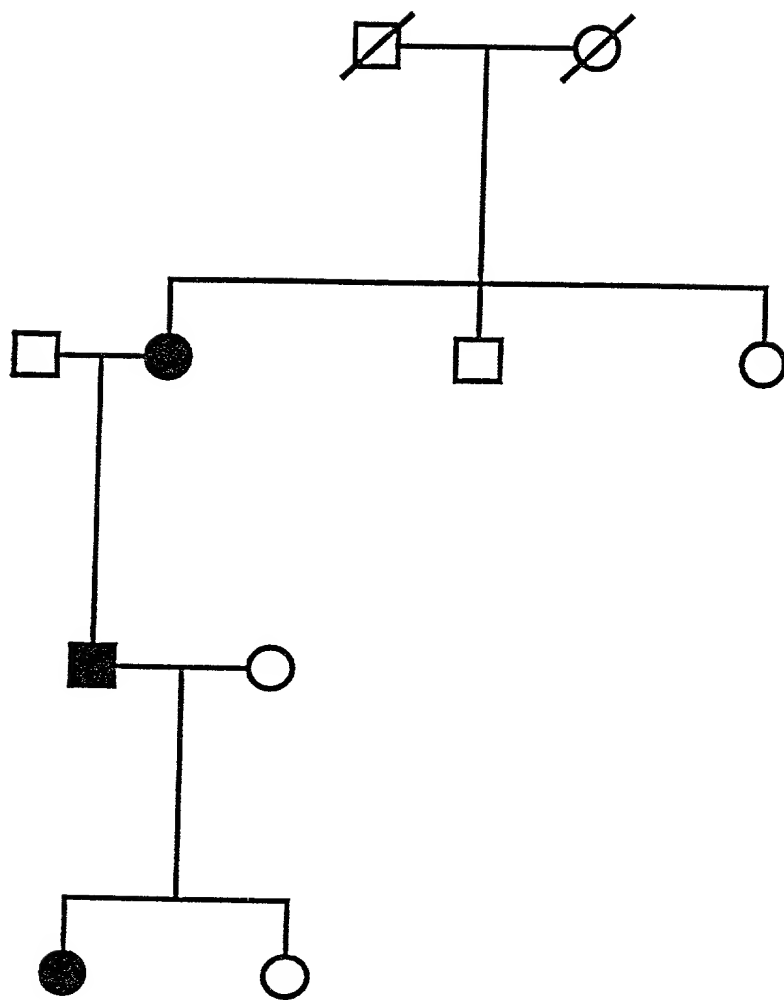


Fig. 1F

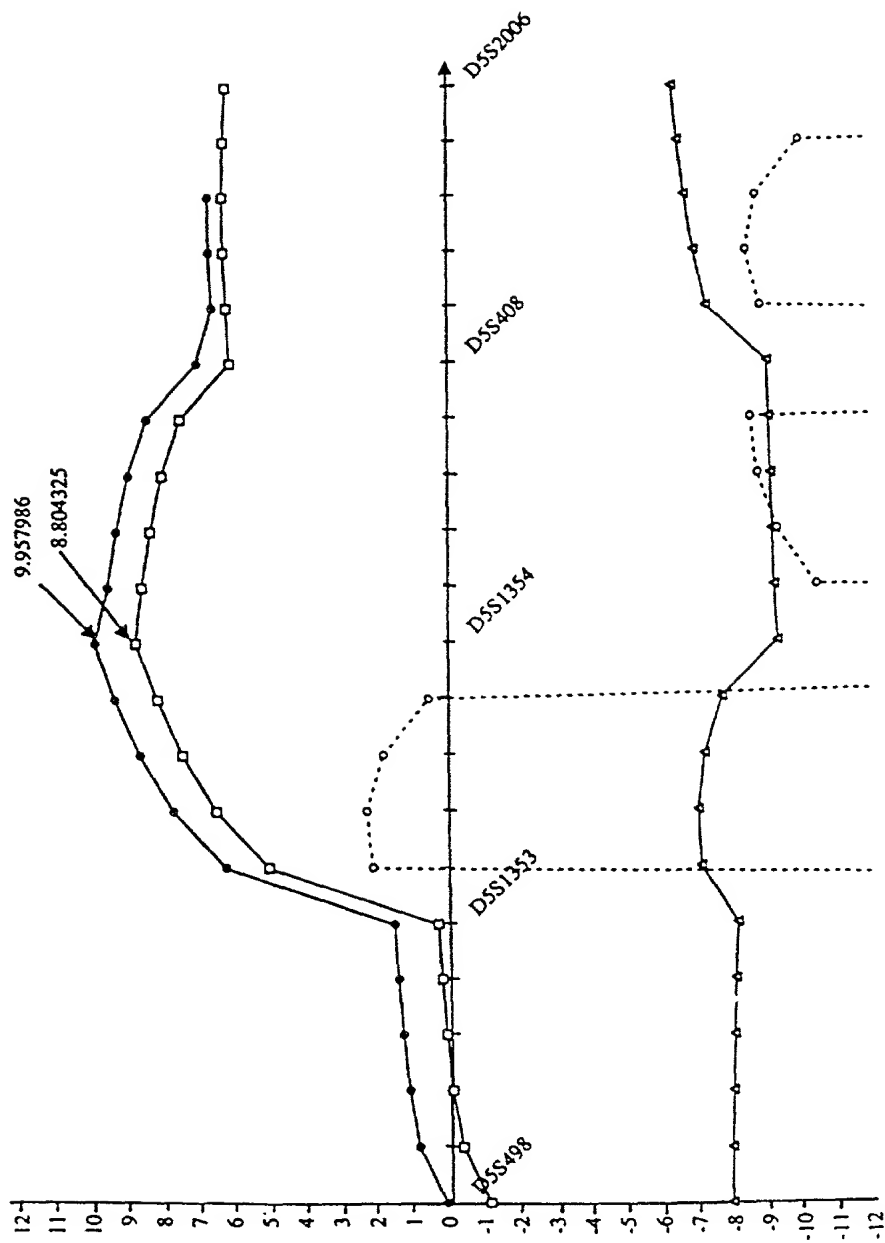


Fig. 2

Species	Accession	Position	Sequence	Length
Human	1		MQRGAALCLRLWLCLGLLDGLVSGYSMTPTNLNITEESHVIDTGDLSLSIS	50
Mouse	1		MQPGAALNLRLWLCLGLLQGLANGYSMTPTNLNITEDSYVIDTGDLSLSIS	50
H	51		CRGQHPLEWAWPGAQEAPATGDKDSED TG VVRDCEGTDARPYCKVLLLHE	100
M	51		CRGQHPLEWTWPGAQEVLT TG GKDSED TR VVHDCEGTEARPYCKVLLLAQ	100
H	101		VHANDTGSYVCYYKYIKARIEGTTAASSYVFVRDFEQPFINKPDTLLVNR	150
M	101		THANNTGSYHCYYKYIKARIEGTTAASTYVFVRDFKHPFINKPDTLLVNR	150
H	151		KDAMWVPCLV SIPGLNVT LRSQSSVLWPDGQEVVWDDRRGMLVSTPLLHD	200
M	151		KDSMWVPCLV SIPGLNIT LRSQSSALHPDGQEV LWDDRRGMRVPTQLLRD	200
H	201		ALYLQCETT WGDQDFLSNPFLVHITGNELYDIQLLPKRSLELLVGEKLV	250
M	201		ALYLQCETT WGDQNFLSNLFVVHITGNELYDIQLYPKKSME LLVGEKLV	250
H	251		NCTVWAEFNSGVTFDWDYPGKQAERGKWPERRSQQTHTELSSILTIHN	300
M	251		NCTVWAEFDSGVTFDWDYPGKQAERAKWVPERRSQQTHTELSSILTIHN	300
H	301		SQHDLSGYVCKANNGIQRFRESTEVI VHENPFISVEWLKGP ILEATAGDE	350
M	301		SQNDLGPYVCEANNGIQRFRESTEVI VHEKPFISVEWLKGPVLEATAGDE	350
H	351		LVKLPVKLAAYPPPEFQWKDGKALSGRHS PHALVLKEVTEASTGTYTLA	400
M	351		LVKLPVKLAAYPPPEFQWKDRKAVTGRHNPHALVLKEVTEASAGVYTLA	400
H	401		LWNSAAGLR RNISLELVNVVPPQIHEKEASSPSIYSRHSRQALTCTAYGV	450
M	401		LWNSAAGLRQNISLELVNVVPPHIHEKEASSPSIYSRHSRQTLTCTAYGV	450
H	451		PLPLSIQWHWRPWP TPKMFAQRSLRRRQQQDLMPQCRDWRAVTTQDAVNP	500
M	451		PQPLSVQWHWRPWP TPKTFAQRSLRRRQQRDGMPQCRDWKEVTTQDAVNP	500
H	501		IESLDTWTEFVEGKNKT VSKLVIQANANVSAMYKCVVSNKVGQDERLIYFY	550
M	501		IESLD SWTEFVEGKNKT VSKLVIQ DANVSAMYKCVVSNKVGQDERLIYFY	550
H	551		VTTIPDGFTIESK PSEELLEGPVLLSCQADSYKYEHLRWYRLNLSTLHD	600
M	551		VTTIPDGFSIESEPS EDPLEGQSVRLSCRADNYTYEHLRWYRLNLSTLHD	600
H	601		AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLSIPRVAP EHEGH	650
M	601		AQGNPLLLDCKNVHLFATPLEANLEEAEPGARHATLSLNIPRVAP EDEGD	650
H	651		YVCEVQDRRSHDKHCHKKYL SVQALEAPRLTQNLTDLLVNVSDSLEMQCL	700
M	651		YVCEVODRRSODKHCHKKYL SVQALEAPRLTQNLTDLLVNVSDSLEMRCP	700

H 701 VAGAHAPSIVWYKDERLLEEKSGVDLADSNQKLSIQRVREEDAGRYLCSV 750
 ||||| ||||| ||||| ||||| . . : ||||| : ||||| ||||| |||||
 M 701 VAGAHVPSIVWYKDERLLEKESGIDLADSNQRLSIQRVREEDAGRYLCSV 750
 H 751 CNAKGCVNSSASVAVEGSEDKGSMEIVILVGTGVIAVFFWVLLLLLIFCNM 800
 ||||| ||||| ||||| ||||| ||||| ||||| : ||||| ||||| ||||| ||||| ||||| |||||
 M 751 CNAKGCVNSSASVAVEGSEDKGSMEIVILIGTGVIAVFFWVLLLLLIFCNM 800
 H 801 RRPAAHADIKTGYSIIIMDPGEVPLEEQCEYLSYDASQWEFPRERLHLGRV 850
 : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 M 801 KRPAHADIKTGYSIIIMDPGEVPLEEQCEYLSYDASQWEFPRERLHLGRV 850
 H 851 LGYGAFGKVVEASAFGIHKGSSCDTVAVKMLKEGATASEHRALMSELKIL 900
 ||||| ||||| ||||| ||||| . . : ||||| ||||| ||||| ||||| ||||| ||||| |||||
 M 851 LGHGAFGKVVEASAFGINKGSSCDTVAVKMLKEGATASEHRALMSELKIL 900
 G857R
 H 901 IHIGNHLNVVNLLGACTKPQGPLMVIVEFCKYGNLSNFLRAKRDAFSPCA 950
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 M 901 IHIGNHLNVVNLLGACTKPNGPLMVIVEFCKYGNLSNFLRVKRDFTFNPYA 950
 H 951 EKSPEQRGRFRAMVELARLDRRRPGSSDRVLFARFSKTEGGARRASPDQE 1000
 ||||| ||||| ||||| ||||| : ||||| ||||| ||||| ||||| . ||||| ||||| |||||
 M 951 EKSPEQRRRFRAMVEGAKADRRRPGSSDRALFTRFLMGKGSARRAPLVQE 1000
 H 1001 AEDLWLSPLTMEDLVCYSFQVARGMEFLASRKCIHRDLAARNILLSESDV 1050
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 M 1001 AEDLWLSPLTMEDLVCYSFQVARGMEFLASRKCIHRDLAARNILLSESDI 1050
 R1041P D1049N
 L1044P
 H 1051 VKICDFGLARDIYKDPDYVRKGSARLPLKWMAPESIFDKVYTTQSDVWSF 1100
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 M 1051 VKICDFGLARDIYKDPDYVRKGSARLPLKWMAPESIFDKVYTTQSDVWSF 1100
 I1053N (Chy mouse mutation)
 H 1101 GVLLWEIFSLGASPYPGVQINEEFCQRLRDGTRMRAPELATPAIRRIMLN 1150
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : ||||| ||||| ||||| |||||
 M 1101 GVLLWEIFSLGASPYPGVQINEEFCQRLKDGTMRAPLATPAIRHIMQS 1150
 P1114L
 H 1151 CWSGDPKARPAFSELVEILGDLLQGRGLQEEEEVCMAPRSSQSSEEGSFS 1200
 ||||| ||||| ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 M 1151 CWSGDPKARPAFSDLVEILGDLLQGGGWQEEEEERMALHSSQSSEEDGFM 1200
 H 1201 QVSTMALHIAQADAEDSPPSLQRHSLAARYYNWVSFPGCLARGAETRGSS 1250
 ||||| ||||| : ||||| : ||||| : ||||| ||||| ||||| ||||| ||||| |||||
 M 1201 QASTTALHITEADADDSPSMHCHSLAARYYNVCVSFPGRLARGTKTPGSS 1250
 H 1251 RMKTFFEEFPMTPTTYKGSVDNQTDSGMVLASEEFEQIESRHRQESGF 1297
 ||||| ||||| ||||| ||||| . ||||| ||||| ||||| ||||| : ||||| ||||| |||||
 M 1251 RMKTFEELPMTPTTYKASMDNQTDSGMVLASEEFEELSRHRPEGSF 1297

FIG. 3B

SEQUENCE LISTING

<110> Ferrell, Robert E.
Alitalo, Kari
Finegold, David N
Karkkainen, Marika

<120> SCREENING AND THERAPY FOR LYMPHATIC DISORDERS INVOLVING
THE FLT4 RECEPTOR TYROSINE KINASE (VEGFR-3)

<130> 28967/35255A

<140>

<141>

<150> PCT/US99/06133

<151> 1999-03-26

<160> 28

<170> PatentIn Ver. 2.0

<210> 1

<211> 4111

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (20)..(4111)

<220>

<223> Human Flt4 (VEGFR-3) long form cDNA

<400> 1

```

ccacgcgcag cggccggag atg cag cgg ggc gcc gcg ctg tgc ctg cga ctg   52
          Met Gln Arg Gly Ala Ala Leu Cys Leu Arg Leu
              1              5              10

tgg ctc tgc ctg gga ctc ctg gac ggc ctg gtg agt ggc tac tcc atg   100
Trp Leu Cys Leu Gly Leu Leu Asp Gly Leu Val Ser Gly Tyr Ser Met
          15              20              25

acc ccc ccg acc ttg aac atc acg gag gag tca cac gtc atc gac acc   148
Thr Pro Pro Thr Leu Asn Ile Thr Glu Glu Ser His Val Ile Asp Thr
          30              35              40

ggt gac agc ctg tcc atc tcc tgc agg gga cag cac ccc ctc gag tgg   196
Gly Asp Ser Leu Ser Ile Ser Cys Arg Gly Gln His Pro Leu Glu Trp
          45              50              55

gct tgg cca gga gct cag gag gcg cca gcc acc gga gac aag gac agc   244
Ala Trp Pro Gly Ala Gln Glu Ala Pro Ala Thr Gly Asp Lys Asp Ser
          60              65              70              75

```


cag aat gcc aac gtg tct gcc atg tac aag tgt gtg gtc tcc aac aag	1636
Gln Asn Ala Asn Val Ser Ala Met Tyr Lys Cys Val Val Ser Asn Lys	
525 530 535	
gtg ggc cag gat gag cgg ctc atc tac ttc tat gtg acc acc atc ccc	1684
Val Gly Gln Asp Glu Arg Leu Ile Tyr Phe Tyr Val Thr Thr Ile Pro	
540 545 550 555	
gac ggc ttc acc atc gaa tcc aag cca tcc gag gag cta cta gag ggc	1732
Asp Gly Phe Thr Ile Glu Ser Lys Pro Ser Glu Glu Leu Leu Glu Gly	
560 565 570	
cag ccg gtg ctc ctg agc tgc caa gcc gac agc tac aag tac gag cat	1780
Gln Pro Val Leu Leu Ser Cys Gln Ala Asp Ser Tyr Lys Tyr Glu His	
575 580 585	
ctg cgc tgg tac cgc ctc aac ctg tcc acg ctg cac gat gcg cac ggg	1828
Leu Arg Trp Tyr Arg Leu Asn Leu Ser Thr Leu His Asp Ala His Gly	
590 595 600	
aac ccg ctt ctg ctc gac tgc aag aac gtg cat ctg ttc gcc acc cct	1876
Asn Pro Leu Leu Leu Asp Cys Lys Asn Val His Leu Phe Ala Thr Pro	
605 610 615	
ctg gcc gcc agc ctg gag gag gtg gca cct ggg gcg cgc cac gcc acg	1924
Leu Ala Ala Ser Leu Glu Glu Val Ala Pro Gly Ala Arg His Ala Thr	
620 625 630 635	
ctc agc ctg agt atc ccc cgc gtc gcg ccc gag cac gag ggc cac tat	1972
Leu Ser Leu Ser Ile Pro Arg Val Ala Pro Glu His Glu Gly His Tyr	
640 645 650	
gtg tgc gaa gtg caa gac cgg cgc agc cat gac aag cac tgc cac aag	2020
Val Cys Glu Val Gln Asp Arg Arg Ser His Asp Lys His Cys His Lys	
655 660 665	
aag tac ctg tcg gtg cag gcc ctg gaa gcc cct cgg ctc acg cag aac	2068
Lys Tyr Leu Ser Val Gln Ala Leu Glu Ala Pro Arg Leu Thr Gln Asn	
670 675 680	
ttg acc gac ctc ctg gtg aac gtg agc gac tcg ctg gag atg cag tgc	2116
Leu Thr Asp Leu Leu Val Asn Val Ser Asp Ser Leu Glu Met Gln Cys	
685 690 695	
ttg gtg gcc gga gcg cac gcg ccc agc atc gtg tgg tac aaa gac gag	2164
Leu Val Ala Gly Ala His Ala Pro Ser Ile Val Trp Tyr Lys Asp Glu	
700 705 710 715	
agg ctg ctg gag gaa aag tct gga gtc gac ttg gcg gac tcc aac cag	2212
Arg Leu Leu Glu Glu Lys Ser Gly Val Asp Leu Ala Asp Ser Asn Gln	
720 725 730	
aag ctg agc atc cag cgc gtg cgc gag gag gat gcg gga cgc tat ctg	2260
Lys Leu Ser Ile Gln Arg Val Arg Glu Glu Asp Ala Gly Arg Tyr Leu	
735 740 745	

tgc agc gtg tgc aac gcc aag ggc tgc gtc aac tcc tcc gcc agc gtg	2308
Cys Ser Val Cys Asn Ala Lys Gly Cys Val Asn Ser Ser Ala Ser Val	
750 755 760	
gcc gtg gaa ggc tcc gag gat aag ggc agc atg gag atc gtg atc ctt	2356
Ala Val Glu Gly Ser Glu Asp Lys Gly Ser Met Glu Ile Val Ile Leu	
765 770 775	
gtc ggt acc ggc gtc atc gct gtc ttc ttc tgg gtc ctc ctc ctc ctc	2404
Val Gly Thr Gly Val Ile Ala Val Phe Phe Trp Val Leu Leu Leu Leu	
780 785 790 795	
atc ttc tgt aac atg agg agg ccg gcc cac gca gac atc aag acg ggc	2452
Ile Phe Cys Asn Met Arg Arg Pro Ala His Ala Asp Ile Lys Thr Gly	
800 805 810	
tac ctg tcc atc atc atg gac ccc ggg gag gtg cct ctg gag gag caa	2500
Tyr Leu Ser Ile Ile Met Asp Pro Gly Glu Val Pro Leu Glu Glu Gln	
815 820 825	
tgc gaa tac ctg tcc tac gat gcc agc cag tgg gaa ttc ccc cga gag	2548
Cys Glu Tyr Leu Ser Tyr Asp Ala Ser Gln Trp Glu Phe Pro Arg Glu	
830 835 840	
cgg ctg cac ctg ggg aga gtg ctc ggc tac ggc gcc ttc ggg aag gtg	2596
Arg Leu His Leu Gly Arg Val Leu Gly Tyr Gly Ala Phe Gly Lys Val	
845 850 855	
gtg gaa gcc tcc gct ttc ggc atc cac aag ggc agc agc tgt gac acc	2644
Val Glu Ala Ser Ala Phe Gly Ile His Lys Gly Ser Ser Cys Asp Thr	
860 865 870 875	
gtg gcc gtg aaa atg ctg aaa gag ggc gcc acg gcc agc gag cac cgc	2692
Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr Ala Ser Glu His Arg	
880 885 890	
gcg ctg atg tcg gag ctc aag atc ctc att cac atc ggc aac cac ctc	2740
Ala Leu Met Ser Glu Leu Lys Ile Leu Ile His Ile Gly Asn His Leu	
895 900 905	
aac gtg gtc aac ctc ctc ggg gcg tgc acc aag ccg cag ggc ccc ctc	2788
Asn Val Val Asn Leu Leu Gly Ala Cys Thr Lys Pro Gln Gly Pro Leu	
910 915 920	
atg gtg atc gtg gag ttc tgc aag tac ggc aac ctc tcc aac ttc ctg	2836
Met Val Ile Val Glu Phe Cys Lys Tyr Gly Asn Leu Ser Asn Phe Leu	
925 930 935	
cgc gcc aag cgg gac gcc ttc agc ccc tgc gcg gag aag tct ccc gag	2884
Arg Ala Lys Arg Asp Ala Phe Ser Pro Cys Ala Glu Lys Ser Pro Glu	
940 945 950 955	
cag cgc gga cgc ttc cgc gcc atg gtg gag ctc gcc agg ctg gat cgg	2932
Gln Arg Gly Arg Phe Arg Ala Met Val Glu Leu Ala Arg Leu Asp Arg	
960 965 970	

agg cgg ccg ggg agc agc gac agg gtc ctc ttc gcg cgg ttc tcg aag	2980
Arg Arg Pro Gly Ser Ser Asp Arg Val Leu Phe Ala Arg Phe Ser Lys	
975 980 985	
acc gag ggc gga gcg agg cgg gct tct cca gac caa gaa gct gag gac	3028
Thr Glu Gly Gly Ala Arg Arg Ala Ser Pro Asp Gln Glu Ala Glu Asp	
990 995 1000	
ctg tgg ctg agc ccg ctg acc atg gaa gat ctt gtc tgc tac agc ttc	3076
Leu Trp Leu Ser Pro Leu Thr Met Glu Asp Leu Val Cys Tyr Ser Phe	
1005 1010 1015	
cag gtg gcc aga ggg atg gag ttc ctg gct tcc cga aag tgc atc cac	3124
Gln Val Ala Arg Gly Met Glu Phe Leu Ala Ser Arg Lys Cys Ile His	
1020 1025 1030 1035	
aga gac ctg gct gct cgg aac att ctg ctg tcg gaa agc gac gtg gtg	3172
Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser Glu Ser Asp Val Val	
1040 1045 1050	
aag atc tgt gac ttt ggc ctt gcc cgg gac atc tac aaa gac cct gac	3220
Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys Asp Pro Asp	
1055 1060 1065	
tac gtc cgc aag ggc agt gcc cgg ctg ccc ctg aag tgg atg gcc cct	3268
Tyr Val Arg Lys Gly Ser Ala Arg Leu Pro Leu Lys Trp Met Ala Pro	
1070 1075 1080	
gaa agc atc ttc gac aag gtg tac acc acg cag agt gac gtg tgg tcc	3316
Glu Ser Ile Phe Asp Lys Val Tyr Thr Thr Gln Ser Asp Val Trp Ser	
1085 1090 1095	
ttt ggg gtg ctt ctc tgg gag atc ttc tct ctg ggg gcc tcc ccg tac	3364
Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu Gly Ala Ser Pro Tyr	
1100 1105 1110 1115	
cct ggg gtg cag atc aat gag gag ttc tgc cag cgg ctg aga gac ggc	3412
Pro Gly Val Gln Ile Asn Glu Glu Phe Cys Gln Arg Leu Arg Asp Gly	
1120 1125 1130	
aca agg atg agg gcc ccg gag ctg gcc act ccc gcc ata cgc cgc atc	3460
Thr Arg Met Arg Ala Pro Glu Leu Ala Thr Pro Ala Ile Arg Arg Ile	
1135 1140 1145	
atg ctg aac tgc tgg tcc gga gac ccc aag gcg aga cct gca ttc tcg	3508
Met Leu Asn Cys Trp Ser Gly Asp Pro Lys Ala Arg Pro Ala Phe Ser	
1150 1155 1160	
gag ctg gtg gag atc ctg ggg gac ctg ctc cag ggc agg ggc ctg caa	3556
Glu Leu Val Glu Ile Leu Gly Asp Leu Leu Gln Gly Arg Gly Leu Gln	
1165 1170 1175	
gag gaa gag gag gtc tgc atg gcc ccg cgc agc tct cag agc tca gaa	3604
Glu Glu Glu Glu Val Cys Met Ala Pro Arg Ser Ser Gln Ser Ser Glu	
1180 1185 1190 1195	

gag	ggc	agc	ttc	tgc	cag	gtg	tcc	acc	atg	gcc	cta	cac	atc	gcc	cag	3652
Glu	Gly	Ser	Phe	Ser	Gln	Val	Ser	Thr	Met	Ala	Leu	His	Ile	Ala	Gln	
			1200						1205					1210		
gct	gac	gct	gag	gac	agc	ccg	cca	agc	ctg	cag	cgc	cac	agc	ctg	gcc	3700
Ala	Asp	Ala	Glu	Asp	Ser	Pro	Pro	Ser	Leu	Gln	Arg	His	Ser	Leu	Ala	
			1215						1220					1225		
gcc	agg	tat	tac	aac	tgg	gtg	tcc	ttt	ccc	ggg	tgc	ctg	gcc	aga	ggg	3748
Ala	Arg	Tyr	Tyr	Asn	Trp	Val	Ser	Phe	Pro	Gly	Cys	Leu	Ala	Arg	Gly	
			1230					1235					1240			
gct	gag	acc	cgt	ggt	tcc	tcc	agg	atg	aag	aca	ttt	gag	gaa	ttc	ccc	3796
Ala	Glu	Thr	Arg	Gly	Ser	Ser	Arg	Met	Lys	Thr	Phe	Glu	Glu	Phe	Pro	
			1245				1250					1255				
atg	acc	cca	acg	acc	tac	aaa	ggc	tct	gtg	gac	aac	cag	aca	gac	agt	3844
Met	Thr	Pro	Thr	Thr	Tyr	Lys	Gly	Ser	Val	Asp	Asn	Gln	Thr	Asp	Ser	
1260					1265					1270					1275	
ggg	atg	gtg	ctg	gcc	tgc	gag	gag	ttt	gag	cag	ata	gag	agc	agg	cat	3892
Gly	Met	Val	Leu	Ala	Ser	Glu	Glu	Phe	Glu	Gln	Ile	Glu	Ser	Arg	His	
				1280					1285					1290		
aga	caa	gaa	agc	ggc	ttc	agc	tgt	aaa	gga	cct	ggc	cag	aat	gtg	gct	3940
Arg	Gln	Glu	Ser	Gly	Phe	Ser	Cys	Lys	Gly	Pro	Gly	Gln	Asn	Val	Ala	
			1295					1300					1305			
gtg	acc	agg	gca	cac	cct	gac	tcc	caa	ggg	agg	cgg	cgg	cgg	cct	gag	3988
Val	Thr	Arg	Ala	His	Pro	Asp	Ser	Gln	Gly	Arg	Arg	Arg	Arg	Pro	Glu	
			1310				1315					1320				
cgg	ggg	gcc	cga	gga	ggc	cag	gtg	ttt	tac	aac	agc	gag	tat	ggg	gag	4036
Arg	Gly	Ala	Arg	Gly	Gly	Gln	Val	Phe	Tyr	Asn	Ser	Glu	Tyr	Gly	Glu	
		1325				1330					1335					
ctg	tgc	gag	cca	agc	gag	gag	gac	cac	tgc	tcc	ccg	tct	gcc	cgc	gtg	4084
Leu	Ser	Glu	Pro	Ser	Glu	Glu	Asp	His	Cys	Ser	Pro	Ser	Ala	Arg	Val	
1340					1345					1350				1355		
act	ttc	ttc	aca	gac	aac	agc	tac	taa								4111
Thr	Phe	Phe	Thr	Asp	Asn	Ser	Tyr									
			1360													
<210>	2															
<211>	1363															
<212>	PRT															
<213>	Homo sapiens															
<400>	2															
Met	Gln	Arg	Gly	Ala	Ala	Leu	Cys	Leu	Arg	Leu	Trp	Leu	Cys	Leu	Gly	
1				5					10					15		
Leu	Leu	Asp	Gly	Leu	Val	Ser	Gly	Tyr	Ser	Met	Thr	Pro	Pro	Thr	Leu	

Asn	Ile	Thr	Glu	Glu	Ser	His	Val	Ile	Asp	Thr	Gly	Asp	Ser	Leu	Ser	
		35					40					45				
Ile	Ser	Cys	Arg	Gly	Gln	His	Pro	Leu	Glu	Trp	Ala	Trp	Pro	Gly	Ala	
	50					55					60					
Gln	Glu	Ala	Pro	Ala	Thr	Gly	Asp	Lys	Asp	Ser	Glu	Asp	Thr	Gly	Val	
65					70					75					80	
Val	Arg	Asp	Cys	Glu	Gly	Thr	Asp	Ala	Arg	Pro	Tyr	Cys	Lys	Val	Leu	
				85					90					95		
Leu	Leu	His	Glu	Val	His	Ala	Asn	Asp	Thr	Gly	Ser	Tyr	Val	Cys	Tyr	
			100					105					110			
Tyr	Lys	Tyr	Ile	Lys	Ala	Arg	Ile	Glu	Gly	Thr	Thr	Ala	Ala	Ser	Ser	
		115					120					125				
Tyr	Val	Phe	Val	Arg	Asp	Phe	Glu	Gln	Pro	Phe	Ile	Asn	Lys	Pro	Asp	
	130					135					140					
Thr	Leu	Leu	Val	Asn	Arg	Lys	Asp	Ala	Met	Trp	Val	Pro	Cys	Leu	Val	
145				150						155					160	
Ser	Ile	Pro	Gly	Leu	Asn	Val	Thr	Leu	Arg	Ser	Gln	Ser	Ser	Val	Leu	
			165						170					175		
Trp	Pro	Asp	Gly	Gln	Glu	Val	Val	Trp	Asp	Asp	Arg	Arg	Gly	Met	Leu	
			180					185					190			
Val	Ser	Thr	Pro	Leu	Leu	His	Asp	Ala	Leu	Tyr	Leu	Gln	Cys	Glu	Thr	
		195					200					205				
Thr	Trp	Gly	Asp	Gln	Asp	Phe	Leu	Ser	Asn	Pro	Phe	Leu	Val	His	Ile	
	210					215					220					
Thr	Gly	Asn	Glu	Leu	Tyr	Asp	Ile	Gln	Leu	Leu	Pro	Arg	Lys	Ser	Leu	
225				230						235					240	
Glu	Leu	Leu	Val	Gly	Glu	Lys	Leu	Val	Leu	Asn	Cys	Thr	Val	Trp	Ala	
			245						250				255			
Glu	Phe	Asn	Ser	Gly	Val	Thr	Phe	Asp	Trp	Asp	Tyr	Pro	Gly	Lys	Gln	
		260						265					270			
Ala	Glu	Arg	Gly	Lys	Trp	Val	Pro	Glu	Arg	Arg	Ser	Gln	Gln	Thr	His	
		275					280					285				
Thr	Glu	Leu	Ser	Ser	Ile	Leu	Thr	Ile	His	Asn	Val	Ser	Gln	His	Asp	
	290					295					300					
Leu	Gly	Ser	Tyr	Val	Cys	Lys	Ala	Asn	Asn	Gly	Ile	Gln	Arg	Phe	Arg	
305				310						315					320	

Glu	Ser	Thr	Glu	Val	Ile	Val	His	Glu	Asn	Pro	Phe	Ile	Ser	Val	Glu		
			325						330						335		
Trp	Leu	Lys	Gly	Pro	Ile	Leu	Glu	Ala	Thr	Ala	Gly	Asp	Glu	Leu	Val		
			340						345						350		
Lys	Leu	Pro	Val	Lys	Leu	Ala	Ala	Tyr	Pro	Pro	Pro	Glu	Phe	Gln	Trp		
			355						360						365		
Tyr	Lys	Asp	Gly	Lys	Ala	Leu	Ser	Gly	Arg	His	Ser	Pro	His	Ala	Leu		
			370						375						380		
Val	Leu	Lys	Glu	Val	Thr	Glu	Ala	Ser	Thr	Gly	Thr	Tyr	Thr	Leu	Ala		
385						390						395			400		
Leu	Trp	Asn	Ser	Ala	Ala	Gly	Leu	Arg	Arg	Asn	Ile	Ser	Leu	Glu	Leu		
			405						410						415		
Val	Val	Asn	Val	Pro	Pro	Gln	Ile	His	Glu	Lys	Glu	Ala	Ser	Ser	Pro		
			420						425						430		
Ser	Ile	Tyr	Ser	Arg	His	Ser	Arg	Gln	Ala	Leu	Thr	Cys	Thr	Ala	Tyr		
			435						440						445		
Gly	Val	Pro	Leu	Pro	Leu	Ser	Ile	Gln	Trp	His	Trp	Arg	Pro	Trp	Thr		
450						455						460					
Pro	Cys	Lys	Met	Phe	Ala	Gln	Arg	Ser	Leu	Arg	Arg	Arg	Gln	Gln	Gln		
465						470						475			480		
Asp	Leu	Met	Pro	Gln	Cys	Arg	Asp	Trp	Arg	Ala	Val	Thr	Thr	Gln	Asp		
			485						490						495		
Ala	Val	Asn	Pro	Ile	Glu	Ser	Leu	Asp	Thr	Trp	Thr	Glu	Phe	Val	Glu		
			500						505						510		
Gly	Lys	Asn	Lys	Thr	Val	Ser	Lys	Leu	Val	Ile	Gln	Asn	Ala	Asn	Val		
			515						520						525		
Ser	Ala	Met	Tyr	Lys	Cys	Val	Val	Ser	Asn	Lys	Val	Gly	Gln	Asp	Glu		
530						535						540					
Arg	Leu	Ile	Tyr	Phe	Tyr	Val	Thr	Thr	Ile	Pro	Asp	Gly	Phe	Thr	Ile		
545						550						555			560		
Glu	Ser	Lys	Pro	Ser	Glu	Glu	Leu	Leu	Glu	Gly	Gln	Pro	Val	Leu	Leu		
			565						570						575		
Ser	Cys	Gln	Ala	Asp	Ser	Tyr	Lys	Tyr	Glu	His	Leu	Arg	Trp	Tyr	Arg		
			580						585						590		
Leu	Asn	Leu	Ser	Thr	Leu	His	Asp	Ala	His	Gly	Asn	Pro	Leu	Leu	Leu		
595						600						605					

Asp	Cys	Lys	Asn	Val	His	Leu	Phe	Ala	Thr	Pro	Leu	Ala	Ala	Ser	Leu	610	615	620	
Glu	Glu	Val	Ala	Pro	Gly	Ala	Arg	His	Ala	Thr	Leu	Ser	Leu	Ser	Ile	625	630	635	640
Pro	Arg	Val	Ala	Pro	Glu	His	Glu	Gly	His	Tyr	Val	Cys	Glu	Val	Gln	645	650	655	
Asp	Arg	Arg	Ser	His	Asp	Lys	His	Cys	His	Lys	Lys	Tyr	Leu	Ser	Val	660	665	670	
Gln	Ala	Leu	Glu	Ala	Pro	Arg	Leu	Thr	Gln	Asn	Leu	Thr	Asp	Leu	Leu	675	680	685	
Val	Asn	Val	Ser	Asp	Ser	Leu	Glu	Met	Gln	Cys	Leu	Val	Ala	Gly	Ala	690	695	700	
His	Ala	Pro	Ser	Ile	Val	Trp	Tyr	Lys	Asp	Glu	Arg	Leu	Leu	Glu	Glu	705	710	715	720
Lys	Ser	Gly	Val	Asp	Leu	Ala	Asp	Ser	Asn	Gln	Lys	Leu	Ser	Ile	Gln	725	730	735	
Arg	Val	Arg	Glu	Glu	Asp	Ala	Gly	Arg	Tyr	Leu	Cys	Ser	Val	Cys	Asn	740	745	750	
Ala	Lys	Gly	Cys	Val	Asn	Ser	Ser	Ala	Ser	Val	Ala	Val	Glu	Gly	Ser	755	760	765	
Glu	Asp	Lys	Gly	Ser	Met	Glu	Ile	Val	Ile	Leu	Val	Gly	Thr	Gly	Val	770	775	780	
Ile	Ala	Val	Phe	Phe	Trp	Val	Leu	Leu	Leu	Leu	Ile	Phe	Cys	Asn	Met	785	790	795	800
Arg	Arg	Pro	Ala	His	Ala	Asp	Ile	Lys	Thr	Gly	Tyr	Leu	Ser	Ile	Ile	805	810	815	
Met	Asp	Pro	Gly	Glu	Val	Pro	Leu	Glu	Glu	Gln	Cys	Glu	Tyr	Leu	Ser	820	825	830	
Tyr	Asp	Ala	Ser	Gln	Trp	Glu	Phe	Pro	Arg	Glu	Arg	Leu	His	Leu	Gly	835	840	845	
Arg	Val	Leu	Gly	Tyr	Gly	Ala	Phe	Gly	Lys	Val	Val	Glu	Ala	Ser	Ala	850	855	860	
Phe	Gly	Ile	His	Lys	Gly	Ser	Ser	Cys	Asp	Thr	Val	Ala	Val	Lys	Met	865	870	875	880
Leu	Lys	Glu	Gly	Ala	Thr	Ala	Ser	Glu	His	Arg	Ala	Leu	Met	Ser	Glu	885	890	895	

Leu Lys Ile Leu Ile His Ile Gly Asn His Leu Asn Val Val Asn Leu
900 905 910

Leu Gly Ala Cys Thr Lys Pro Gln Gly Pro Leu Met Val Ile Val Glu
915 920 925

Phe Cys Lys Tyr Gly Asn Leu Ser Asn Phe Leu Arg Ala Lys Arg Asp
930 935 940

Ala Phe Ser Pro Cys Ala Glu Lys Ser Pro Glu Gln Arg Gly Arg Phe
945 950 955 960

Arg Ala Met Val Glu Leu Ala Arg Leu Asp Arg Arg Arg Pro Gly Ser
965 970 975

Ser Asp Arg Val Leu Phe Ala Arg Phe Ser Lys Thr Glu Gly Gly Ala
980 985 990

Arg Arg Ala Ser Pro Asp Gln Glu Ala Glu Asp Leu Trp Leu Ser Pro
995 1000 1005

Leu Thr Met Glu Asp Leu Val Cys Tyr Ser Phe Gln Val Ala Arg Gly
1010 1015 1020

Met Glu Phe Leu Ala Ser Arg Lys Cys Ile His Arg Asp Leu Ala Ala
1025 1030 1035 1040

Arg Asn Ile Leu Leu Ser Glu Ser Asp Val Val Lys Ile Cys Asp Phe
1045 1050 1055

Gly Leu Ala Arg Asp Ile Tyr Lys Asp Pro Asp Tyr Val Arg Lys Gly
1060 1065 1070

Ser Ala Arg Leu Pro Leu Lys Trp Met Ala Pro Glu Ser Ile Phe Asp
1075 1080 1085

Lys Val Tyr Thr Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu
1090 1095 1100

Trp Glu Ile Phe Ser Leu Gly Ala Ser Pro Tyr Pro Gly Val Gln Ile
1105 1110 1115 1120

Asn Glu Glu Phe Cys Gln Arg Leu Arg Asp Gly Thr Arg Met Arg Ala
1125 1130 1135

Pro Glu Leu Ala Thr Pro Ala Ile Arg Arg Ile Met Leu Asn Cys Trp
1140 1145 1150

Ser Gly Asp Pro Lys Ala Arg Pro Ala Phe Ser Glu Leu Val Glu Ile
1155 1160 1165

Leu Gly Asp Leu Leu Gln Gly Arg Gly Leu Gln Glu Glu Glu Val
1170 1175 1180

Cys Met Ala Pro Arg Ser Ser Gln Ser Ser Glu Glu Gly Ser Phe Ser
 185 1190 1195 1200

Gln Val Ser Thr Met Ala Leu His Ile Ala Gln Ala Asp Ala Glu Asp
 1205 1210 1215

Ser Pro Pro Ser Leu Gln Arg His Ser Leu Ala Ala Arg Tyr Tyr Asn
 1220 1225 1230

Trp Val Ser Phe Pro Gly Cys Leu Ala Arg Gly Ala Glu Thr Arg Gly
 1235 1240 1245

Ser Ser Arg Met Lys Thr Phe Glu Glu Phe Pro Met Thr Pro Thr Thr
 1250 1255 1260

Tyr Lys Gly Ser Val Asp Asn Gln Thr Asp Ser Gly Met Val Leu Ala
 1265 1270 1275 1280

Ser Glu Glu Phe Glu Gln Ile Glu Ser Arg His Arg Gln Glu Ser Gly
 1285 1290 1295

Phe Ser Cys Lys Gly Pro Gly Gln Asn Val Ala Val Thr Arg Ala His
 1300 1305 1310

Pro Asp Ser Gln Gly Arg Arg Arg Arg Pro Glu Arg Gly Ala Arg Gly
 1315 1320 1325

Gly Gln Val Phe Tyr Asn Ser Glu Tyr Gly Glu Leu Ser Glu Pro Ser
 1330 1335 1340

Glu Glu Asp His Cys Ser Pro Ser Ala Arg Val Thr Phe Phe Thr Asp
 1345 1350 1355 1360

Asn Ser Tyr

<210> 3

<211> 1997

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (352)..(1611)

<220>

<223> Human cDNA for prepro-VEGF-C

<400> 3

cccgcgccgc ctctccaaaa agctacaccg acgcggaccg cggcggcgctc ctccctcgcc 60

ctcgcttcac ctgcggggct ccgaatgcgg ggagctcgga tgtccgggtt cctgtgaggc 120

ttttacctga caccgcgcgc ctttccccgg cactggctgg gagggcgccc tgcaaagttg 180

ggaacgcgga gccccggacc cgctcccgcc gcctccggct cgcccagggg gggctcgccgg 240

gaggagcccc ggggagaggg accaggaggg gcccgcgccc tcgcaggggc gcccgcgccc 300

ccacccctgc ccccgccagc ggaccgggtcc cccacccccg gtccttccac c atg cac 357
Met His
1

ttg ctg ggc ttc ttc tct gtg gcg tgt tct ctg ctc gcc gct gcg ctg 405
Leu Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala Ala Leu
5 10 15

ctc ccg ggt cct cgc gag gcg ccc gcc gcc gcc gcc gcc ttc gag tcc 453
Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Ala Phe Glu Ser
20 25 30

gga ctc gac ctc tcg gac gcg gag ccc gac gcg ggc gag gcc acg gct 501
Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala Thr Ala
35 40 45 50

tat gca agc aaa gat ctg gag gag cag tta cgg tct gtg tcc agt gta 549
Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser Ser Val
55 60 65

gat gaa ctc atg act gta ctc tac cca gaa tat tgg aaa atg tac aag 597
Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met Tyr Lys
70 75 80

tgt cag cta agg aaa gga ggc tgg caa cat aac aga gaa cag gcc aac 645
Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln Ala Asn
85 90 95

ctc aac tca agg aca gaa gag act ata aaa ttt gct gca gca cat tat 693
Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr
100 105 110

aat aca gag atc ttg aaa agt att gat aat gag tgg aga aag act caa 741
Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln
115 120 125 130

tgc atg cca cgg gag gtg tgt ata gat gtg ggg aag gag ttt gga gtc 789
Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe Gly Val
135 140 145

gcg aca aac acc ttc ttt aaa cct cca tgt gtg tcc gtc tac aga tgt 837
Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys
150 155 160

ggg ggt tgc tgc aat agt gag ggg ctg cag tgc atg aac acc agc acg 885
Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr Ser Thr
165 170 175

agc tac ctc agc aag acg tta ttt gaa att aca gtg cct ctc tct caa 933
Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln
180 185 190

CCAGGCGCC

ggc ccc aaa cca gta aca atc agt ttt gcc aat cac act tcc tgc cga	981
Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg	
195 200 205 210	
tgc atg tct aaa ctg gat gtt tac aga caa gtt cat tcc att att aga	1029
Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg	
215 220 225	
cgt tcc ctg cca gca aca cta cca cag tgt cag gca gcg aac aag acc	1077
Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr	
230 235 240	
tgc ccc acc aat tac atg tgg aat aat cac atc tgc aga tgc ctg gct	1125
Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys Leu Ala	
245 250 255	
cag gaa gat ttt atg ttt tcc tgc gat gct gga gat gac tca aca gat	1173
Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser Thr Asp	
260 265 270	
gga ttc cat gac atc tgt gga cca aac aag gag ctg gat gaa gag acc	1221
Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu Glu Thr	
275 280 285 290	
tgt cag tgt gtc tgc aga gcg ggg ctt cgg cct gcc agc tgt gga ccc	1269
Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys Gly Pro	
295 300 305	
cac aaa gaa cta gac aga aac tca tgc cag tgt gtc tgt aaa aac aaa	1317
His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys Asn Lys	
310 315 320	
ctc ttc ccc agc caa tgt ggg gcc aac cga gaa ttt gat gaa aac aca	1365
Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr	
325 330 335	
tgc cag tgt gta tgt aaa aga acc tgc ccc aga aat caa ccc cta aat	1413
Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn	
340 345 350	
cct gga aaa tgt gcc tgt gaa tgt aca gaa agt cca cag aaa tgc ttg	1461
Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys Cys Leu	
355 360 365 370	
tta aaa gga aag aag ttc cac cac caa aca tgc agc tgt tac aga cgg	1509
Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg	
375 380 385	
cca tgt acg aac cgc cag aag gct tgt gag cca gga ttt tca tat agt	1557
Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser Tyr Ser	
390 395 400	
gaa gaa gtg tgt cgt tgt gtc cct tca tat tgg aaa aga cca caa atg	1605
Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro Gln Met	
405 410 415	

agc taa gattgtactg ttttccagtt catcgatttt ctattatgga aaactgtgtt 1661
Ser
420

gccacagtag aactgtctgt gaacagagag acccttgtgg gtccatgcta acaaagacaa 1721
aagtctgtct ttcctgaacc atgtggataa ctttacagaa atggactgga gctcatctgc 1781
aaaaggcctc ttgtaaagac tggttttctg ccaatgacca aacagccaag attttctct 1841
tgtgatttct ttaaaagaat gactatataa tttatttcca ctaaaaatat tgtttctgca 1901
ttcattttta tagcaacaac aattggtaaa actcactgtg atcaatattt ttatatcatg 1961
caaaatatgt ttaaaataaa atgaaaattg tattat 1997

<210> 4
<211> 419
<212> PRT
<213> Homo sapiens

<400> 4
Met His Leu Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala
1 5 10 15
Ala Leu Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Phe
20 25 30
Glu Ser Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala
35 40 45
Thr Ala Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser
50 55 60
Ser Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met
65 70 75 80
Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln
85 90 95
Ala Asn Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala
100 105 110
His Tyr Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys
115 120 125
Thr Gln Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe
130 135 140
Gly Val Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr
145 150 155 160
Arg Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr
165 170 175

003343-03400

Ser Thr Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu
180 185 190

Ser Gln Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser
195 200 205

Cys Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile
210 215 220

Ile Arg Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn
225 230 235 240

Lys Thr Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys
245 250 255

Leu Ala Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser
260 265 270

Thr Asp Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu
275 280 285

Glu Thr Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys
290 295 300

Gly Pro His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys
305 310 315 320

Asn Lys Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu
325 330 335

Asn Thr Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro
340 345 350

Leu Asn Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys
355 360 365

Cys Leu Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr
370 375 380

Arg Arg Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser
385 390 395 400

Tyr Ser Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro
405 410 415

Gln Met Ser

<210> 5

<211> 2029

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<223> Human prepro-VEGF-D cDNA

```

gttggggttcc agcttttctgt agctgtaagc attggtggcc acaccacctc cttacaaaagc 60
aactagaacc tgcggcatac attggagaga tttttttaat tttctggaca tgaagtaaata 120
ttagagtgct ttctaatttc aggtagaaga catgtccacc ttctgattat ttttggagaa 180
catttttgatt tttttcatct ctctctcccc acccctaaga ttgtgcaaaa aaagcgtacc 240
ttgcctaatt gaaataattt cattggattt tgatcagaac tgattatttg gttttctgtg 300
tgaagttttg aggtttcaaa ctttccttct ggagaatgcc ttttgaaaca attttctcta 360
gctgcctgat gtcaactgct tagtaatcag tggatattga aatattcaaa atg tac      416
                                     Met Tyr
                                     1

```

ttc	ttc	aag	ccc	cct	tgt	gtg	aac	gtg	ttc	cga	tgt	ggg	ggc	tgt	tgc	848
Phe	Phe	Lys	Pro	Pro	Cys	Val	Asn	Val	Phe	Arg	Cys	Gly	Gly	Cys	Cys	
				135					140				145			

aat gaa gag agc ctt atc tgt atg aac acc agc acc tcg tac att tcc	896
Asn Glu Glu Ser Leu Ile Cys Met Asn Thr Ser Thr Ser Tyr Ile Ser	
150 155 160	
aaa cag ctc ttt gag ata tca gtg cct ttg aca tca gta cct gaa tta	944
Lys Gln Leu Phe Glu Ile Ser Val Pro Leu Thr Ser Val Pro Glu Leu	
165 170 175	
gtg cct gtt aaa gtt gcc aat cat aca ggt tgt aag tgc ttg cca aca	992
Val Pro Val Lys Val Ala Asn His Thr Gly Cys Lys Cys Leu Pro Thr	
180 185 190	
gcc ccc cgc cat cca tac tca att atc aga aga tcc atc cag atc cct	1040
Ala Pro Arg His Pro Tyr Ser Ile Ile Arg Arg Ser Ile Gln Ile Pro	
195 200 205 210	
gaa gaa gat cgc tgt tcc cat tcc aag aaa ctc tgt cct att gac atg	1088
Glu Glu Asp Arg Cys Ser His Ser Lys Lys Leu Cys Pro Ile Asp Met	
215 220 225	
cta tgg gat agc aac aaa tgt aaa tgt gtt ttg cag gag gaa aat cca	1136
Leu Trp Asp Ser Asn Lys Cys Lys Cys Val Leu Gln Glu Glu Asn Pro	
230 235 240	
ctt gct gga aca gaa gac cac tct cat ctc cag gaa cca gct ctc tgt	1184
Leu Ala Gly Thr Glu Asp His Ser His Leu Gln Glu Pro Ala Leu Cys	
245 250 255	
ggg cca cac atg atg ttt gac gaa gat cgt tgc gag tgt gtc tgt aaa	1232
Gly Pro His Met Met Phe Asp Glu Asp Arg Cys Glu Cys Val Cys Lys	
260 265 270	
aca cca tgt ccc aaa gat cta atc cag cac ccc aaa aac tgc agt tgc	1280
Thr Pro Cys Pro Lys Asp Leu Ile Gln His Pro Lys Asn Cys Ser Cys	
275 280 285 290	
ttt gag tgc aaa gaa agt ctg gag acc tgc tgc cag aag cac aag cta	1328
Phe Glu Cys Lys Glu Ser Leu Glu Thr Cys Cys Gln Lys His Lys Leu	
295 300 305	
ttt cac cca gac acc tgc agc tgt gag gac aga tgc ccc ttt cat acc	1376
Phe His Pro Asp Thr Cys Ser Cys Glu Asp Arg Cys Pro Phe His Thr	
310 315 320	
aga cca tgt gca agt ggc aaa aca gca tgt gca aag cat tgc cgc ttt	1424
Arg Pro Cys Ala Ser Gly Lys Thr Ala Cys Ala Lys His Cys Arg Phe	
325 330 335	
cca aag gag aaa agg gct gcc cag ggg ccc cac agc cga aag aat cct	1472
Pro Lys Glu Lys Arg Ala Ala Gln Gly Pro His Ser Arg Lys Asn Pro	
340 345 350	
tga ttcagcgttc caagttcccc atccctgtca tttttaacag catgctgctt	1525

tgccaagttg ctgtcactgt ttttttccca ggtgttaaaa aaaaaatcca ttttacacag 1585
caccacagtg aatccagacc aaccttccat tcacaccagc taaggagtcc ctgggttcatt 1645
gatggatgtc ttctagctgc agatgcctct gcgcaccaag gaatggagag gaggggaccc 1705
atgtaatcct tttgtttagt tttgtttttg ttttttggtg aatgagaaaag gtgtgctggt 1765
catggaatgg caggtgtcat atgactgatt actcagagca gatgaggaaa actgtagtct 1825
ctgagtcctt tgctaatcgc aactcttggtg aattattctg attctttttt atgcagaatt 1885
tgattcgtat gatcagtact gactttctga ttactgtcca gcttatagtc ttccagttta 1945
atgaactacc atctgatgtt tcatatttaa gtgtatttaa agaaaataaa caccattatt 2005
caagccaaaa aaaaaaaaaa aaaa 2029

<210> 6

<211> 354

<212> PRT

<213> Homo sapiens

<400> 6

Met	Tyr	Arg	Glu	Trp	Val	Val	Val	Asn	Val	Phe	Met	Met	Leu	Tyr	Val
1				5					10					15	

Gln	Leu	Val	Gln	Gly	Ser	Ser	Asn	Glu	His	Gly	Pro	Val	Lys	Arg	Ser
		20						25					30		

Ser	Gln	Ser	Thr	Leu	Glu	Arg	Ser	Glu	Gln	Gln	Ile	Arg	Ala	Ala	Ser
		35					40					45			

Ser	Leu	Glu	Glu	Leu	Leu	Arg	Ile	Thr	His	Ser	Glu	Asp	Trp	Lys	Leu
	50					55						60			

Trp	Arg	Cys	Arg	Leu	Arg	Leu	Lys	Ser	Phe	Thr	Ser	Met	Asp	Ser	Arg
65					70					75					80

Ser	Ala	Ser	His	Arg	Ser	Thr	Arg	Phe	Ala	Ala	Thr	Phe	Tyr	Asp	Ile
			85						90					95	

Glu	Thr	Leu	Lys	Val	Ile	Asp	Glu	Glu	Trp	Gln	Arg	Thr	Gln	Cys	Ser
		100						105					110		

Pro	Arg	Glu	Thr	Cys	Val	Glu	Val	Ala	Ser	Glu	Leu	Gly	Lys	Ser	Thr
		115					120					125			

Asn	Thr	Phe	Phe	Lys	Pro	Pro	Cys	Val	Asn	Val	Phe	Arg	Cys	Gly	Gly
	130						135					140			

Cys	Cys	Asn	Glu	Glu	Ser	Leu	Ile	Cys	Met	Asn	Thr	Ser	Thr	Ser	Tyr
145					150					155					160

Ile	Ser	Lys	Gln	Leu	Phe	Glu	Ile	Ser	Val	Pro	Leu	Thr	Ser	Val	Pro
				165					170					175	
Glu	Leu	Val	Pro	Val	Lys	Val	Ala	Asn	His	Thr	Gly	Cys	Lys	Cys	Leu
			180					185					190		
Pro	Thr	Ala	Pro	Arg	His	Pro	Tyr	Ser	Ile	Ile	Arg	Arg	Ser	Ile	Gln
		195					200				205				
Ile	Pro	Glu	Glu	Asp	Arg	Cys	Ser	His	Ser	Lys	Lys	Leu	Cys	Pro	Ile
	210					215				220					
Asp	Met	Leu	Trp	Asp	Ser	Asn	Lys	Cys	Lys	Cys	Val	Leu	Gln	Glu	Glu
225					230					235					240
Asn	Pro	Leu	Ala	Gly	Thr	Glu	Asp	His	Ser	His	Leu	Gln	Glu	Pro	Ala
				245					250					255	
Leu	Cys	Gly	Pro	His	Met	Met	Phe	Asp	Glu	Asp	Arg	Cys	Glu	Cys	Val
			260					265					270		
Cys	Lys	Thr	Pro	Cys	Pro	Lys	Asp	Leu	Ile	Gln	His	Pro	Lys	Asn	Cys
		275					280					285			
Ser	Cys	Phe	Glu	Cys	Lys	Glu	Ser	Leu	Glu	Thr	Cys	Cys	Gln	Lys	His
	290					295					300				
Lys	Leu	Phe	His	Pro	Asp	Thr	Cys	Ser	Cys	Glu	Asp	Arg	Cys	Pro	Phe
305					310					315					320
His	Thr	Arg	Pro	Cys	Ala	Ser	Gly	Lys	Thr	Ala	Cys	Ala	Lys	His	Cys
				325				330						335	
Arg	Phe	Pro	Lys	Glu	Lys	Arg	Ala	Ala	Gln	Gly	Pro	His	Ser	Arg	Lys
			340					345					350		

Asn Pro

<210> 7

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 7

tcaccatcga tccaagc

17

<210> 8

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 8

agttctgcgt gagccgag

18

<210> 9

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 9

caggacgggg tgacttga

18

<210> 10

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 10

gcccaggcct gtctactg

18

<210> 11

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 11

ccagctccta cgtgttcg

18

<210> 12

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 12

ggcaacagct ggatgtca

18

<210> 13
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 13
ctgtgagggc gtgggagt 18

<210> 14
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 14
gtcctttgag ccactgga 18

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 15
cacacgtcat cgacaccggt g 21

<210> 16
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 16
ggcaacagct ggatgtca 18

<210> 17
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:

oligonucleotide

<400> 17
cctgagtatc tcccgcgtcg c 21

<210> 18
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide

<400> 18
ggtgccctccc tgtaccctgg g 21

<210> 19
<211> 1363
<212> PRT
<213> Mus musculus

<400> 19
Met Gln Pro Gly Ala Ala Leu Asn Leu Arg Leu Trp Leu Cys Leu Gly
1 5 10 15
Leu Leu Gln Gly Leu Ala Asn Gly Tyr Ser Met Thr Pro Pro Thr Leu
20 25 30
Asn Ile Thr Glu Asp Ser Tyr Val Ile Asp Thr Gly Asp Ser Leu Ser
35 40 45
Ile Ser Cys Arg Gly Gln His Pro Leu Glu Trp Thr Trp Pro Gly Ala
50 55 60
Gln Glu Val Leu Thr Thr Gly Gly Lys Asp Ser Glu Asp Thr Arg Val
65 70 75 80
Val His Asp Cys Glu Gly Thr Glu Ala Arg Pro Tyr Cys Lys Val Leu
85 90 95
Leu Leu Ala Gln Thr His Ala Asn Asn Thr Gly Ser Tyr His Cys Tyr
100 105 110
Tyr Lys Tyr Ile Lys Ala Arg Ile Glu Gly Thr Thr Ala Ala Ser Thr
115 120 125
Tyr Val Phe Val Arg Asp Phe Lys His Pro Phe Ile Asn Lys Pro Asp
130 135 140
Thr Leu Leu Val Asn Arg Lys Asp Ser Met Trp Val Pro Cys Leu Val
145 150 155 160

Gly	Val	Pro	Gln	Pro	Leu	Ser	Val	Gln	Trp	His	Trp	Arg	Pro	Trp	Thr	450	455	460	
Pro	Cys	Lys	Thr	Phe	Ala	Gln	Arg	Ser	Leu	Arg	Arg	Arg	Gln	Gln	Arg	465	470	475	480
Asp	Gly	Met	Pro	Gln	Cys	Arg	Asp	Trp	Lys	Glu	Val	Thr	Thr	Gln	Asp	485	490	495	
Ala	Val	Asn	Pro	Ile	Glu	Ser	Leu	Asp	Ser	Trp	Thr	Glu	Phe	Val	Glu	500	505	510	
Gly	Lys	Asn	Lys	Thr	Val	Ser	Lys	Leu	Val	Ile	Gln	Asp	Ala	Asn	Val	515	520	525	
Ser	Ala	Met	Tyr	Lys	Cys	Val	Val	Val	Asn	Lys	Val	Gly	Gln	Asp	Glu	530	535	540	
Arg	Leu	Ile	Tyr	Phe	Tyr	Val	Thr	Thr	Ile	Pro	Asp	Gly	Phe	Ser	Ile	545	550	555	560
Glu	Ser	Glu	Pro	Ser	Glu	Asp	Pro	Leu	Glu	Gly	Gln	Ser	Val	Arg	Leu	565	570	575	
Ser	Cys	Arg	Ala	Asp	Asn	Tyr	Thr	Tyr	Glu	His	Leu	Arg	Trp	Tyr	Arg	580	585	590	
Leu	Asn	Leu	Ser	Thr	Leu	His	Asp	Ala	Gln	Gly	Asn	Pro	Leu	Leu	Leu	595	600	605	
Asp	Cys	Lys	Asn	Val	His	Leu	Phe	Ala	Thr	Pro	Leu	Glu	Ala	Asn	Leu	610	615	620	
Glu	Glu	Ala	Glu	Pro	Gly	Ala	Arg	His	Ala	Thr	Leu	Ser	Leu	Asn	Ile	625	630	635	640
Pro	Arg	Val	Ala	Pro	Glu	Asp	Glu	Gly	Asp	Tyr	Val	Cys	Glu	Val	Gln	645	650	655	
Asp	Arg	Arg	Ser	Gln	Asp	Lys	His	Cys	His	Lys	Lys	Tyr	Leu	Ser	Val	660	665	670	
Gln	Ala	Leu	Glu	Ala	Pro	Arg	Leu	Thr	Gln	Asn	Leu	Thr	Asp	Leu	Leu	675	680	685	
Val	Asn	Val	Ser	Asp	Ser	Leu	Glu	Met	Arg	Cys	Pro	Val	Ala	Gly	Ala	690	695	700	
His	Val	Pro	Ser	Ile	Val	Trp	Tyr	Lys	Asp	Glu	Arg	Leu	Leu	Glu	Lys	705	710	715	720
Glu	Ser	Gly	Ile	Asp	Leu	Ala	Asp	Ser	Asn	Gln	Arg	Leu	Ser	Ile	Gln	725	730	735	

Arg	Val	Arg	Glu	Glu	Asp	Ala	Gly	Arg	Tyr	Leu	Cys	Ser	Val	Cys	Asn	
			740						745						750	
Ala	Lys	Gly	Cys	Val	Asn	Ser	Ser	Ala	Ser	Val	Ala	Val	Glu	Gly	Ser	
			755						760						765	
Glu	Asp	Lys	Gly	Ser	Met	Glu	Ile	Val	Ile	Leu	Ile	Gly	Thr	Gly	Val	
			770						775						780	
Ile	Ala	Val	Phe	Phe	Trp	Val	Leu	Leu	Leu	Leu	Ile	Phe	Cys	Asn	Met	
785						790						795			800	
Lys	Arg	Pro	Ala	His	Ala	Asp	Ile	Lys	Thr	Gly	Tyr	Leu	Ser	Ile	Ile	
			805						810						815	
Met	Asp	Pro	Gly	Glu	Val	Pro	Leu	Glu	Glu	Gln	Cys	Glu	Tyr	Leu	Ser	
			820						825						830	
Tyr	Asp	Ala	Ser	Gln	Trp	Glu	Phe	Pro	Arg	Glu	Arg	Leu	His	Leu	Gly	
			835						840						845	
Arg	Val	Leu	Gly	His	Gly	Ala	Phe	Gly	Lys	Val	Val	Glu	Ala	Ser	Ala	
850						855						860				
Phe	Gly	Ile	Asn	Lys	Gly	Ser	Ser	Cys	Asp	Thr	Val	Ala	Val	Lys	Met	
865						870						875			880	
Leu	Lys	Glu	Gly	Ala	Thr	Ala	Ser	Glu	His	Arg	Ala	Leu	Met	Ser	Glu	
			885						890						895	
Leu	Lys	Ile	Leu	Ile	His	Ile	Gly	Asn	His	Leu	Asn	Val	Val	Asn	Leu	
			900						905						910	
Leu	Gly	Ala	Cys	Thr	Lys	Pro	Asn	Gly	Pro	Leu	Met	Val	Ile	Val	Glu	
915						920						925				
Phe	Cys	Lys	Tyr	Gly	Asn	Leu	Ser	Asn	Phe	Leu	Arg	Val	Lys	Arg	Asp	
930						935						940				
Thr	Phe	Asn	Pro	Tyr	Ala	Glu	Lys	Ser	Pro	Glu	Gln	Arg	Arg	Arg	Phe	
945						950						955			960	
Arg	Ala	Met	Val	Glu	Gly	Ala	Lys	Ala	Asp	Arg	Arg	Arg	Pro	Gly	Ser	
			965						970						975	
Ser	Asp	Arg	Ala	Leu	Phe	Thr	Arg	Phe	Leu	Met	Gly	Lys	Gly	Ser	Ala	
			980						985						990	
Arg	Arg	Ala	Pro	Leu	Val	Gln	Glu	Ala	Glu	Asp	Leu	Trp	Leu	Ser	Pro	
995						1000						1005				
Leu	Thr	Met	Glu	Asp	Leu	Val	Cys	Tyr	Ser	Phe	Gln	Val	Ala	Arg	Gly	
1010						1015						1020				

Pro Asp Pro Gln Gly Arg Arg Arg Arg Pro Thr Gln Gly Ala Gln Gly
1315 1320 1325

Gly Lys Val Phe Tyr Asn Asn Glu Tyr Gly Glu Val Ser Gln Pro Cys
1330 1335 1340

Thr Glu Gly Asp Cys Cys Pro Ser Ala Gly Ser Thr Phe Phe Ala Asp
1345 1350 1355 1360

Ser Ser Tyr

<210> 20

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligonucleotide

<400> 20

cggcgccttc aggaaggtgg t

21

<210> 21

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligonucleotide

<400> 21

cggaacattc cgctgtcgga a

21

<210> 22

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligonucleotide

<400> 22

gtcggaaagc aacgtggtga a

21

<210> 23

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 23

catcaagacg ggctacct

18

<210> 24

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 24

ccgctgaccc cacacctt

18

<210> 25

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 25

gagttgacct cccaaggt

18

<210> 26

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 26

tctcctggac aggcagtc

18

<210> 27

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 27

gagttgacct cccaaggt

18

<211> 18

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 28

tctcctggac aggcagtc

18